DOCUMENT RESUME

ED 442 940 CE 080 328

AUTHOR Gorard, Stephen; Rees, Gareth; Furlong, John; Fevre, Ralph

TITLE The Region of the Study [and] Outline Methodology of the Study. Patterns of Participation in Adult Education and Training. Working Paper 1 [and] Working Paper 2.

INSTITUTION Bristol Univ. (England).; Cardiff Univ. (Wales). School of Education.

SPONS AGENCY Economic and Social Research Council, Lancaster (England).
ISBN ISBN-1-872330-03-7; ISBN-1-872330-02-9
PUB DATE 1997-00-00
NOTE 99p.; A Cardiff and Bristol University ESRC-Funded Learning Society Project. Also funded by the Gwent, Mid Glamorgan and West Wales Training Enterprise Councils. For other Working Papers, see CE 080 330-341 and CE 080 402.
CONTRACT ESRC-L123251041
PUB TYPE Reports - Research (143)
EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS *Adult Education; *Attendance Patterns; Developed Nations; Educational Research; Foreign Countries; *Geographic Regions; *Job Training; Participant Characteristics; *Participation; Trend Analysis; *Vocational Education
IDENTIFIERS Learning Patterns; Regional Surveys; *Wales (South)

ABSTRACT

These two working papers are products of a regional study in industrial South Wales of the determinants of participation and non-participation in post-compulsory education and training, with special reference to processes of change in the patterns of these determinants over time and to variations between geographical areas. The study combines contextual analysis of secondary data about education and training providers with a regional study of several generations of families in South Wales via survey, semi-structured interviews, and taped oral histories. Paper 1, The Region of Study, describes the region and provides secondary data used both in the creation of the sample and as a background for the analysis of research findings. It discusses how these three sites that typified the focus area were selected: Blaenau Gwent in the heart of the depressed coalfield; Neath Port Talbot, a large urban center historically linked with heavy industry; and Brigend, a rapidly expanding town south of the coalfield which also shows recent impact of Japanese investment in the area. Paper 1 contains 48 references; background data of the sites; and maps. Paper 2, Outline Methodology of the Study, describes the survey, data collection, pilot study, coding, and these forms of analysis that were carried out: factor analysis, causation, temporal sequences, standard regression, logistic regression, and loglinear models. Paper 2 contains include 63 references. (YLB)

Reproductions supplied by EDRS are the best that can be made from the original document.

ED 442 940

Patterns of Participation in Adult Education and Training

A Cardiff and Bristol University ESRC-funded Learning Society Project

Working Paper 1:

The Region of Study

By Stephen Gorard (1997)

[and]

Working Paper 2:

Outline Methodology of the Study

By Stephen Gorard, Gareth Rees, John Furlong, and Ralph Fevre (1997)

Cardiff University of Wales, School of Education

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

S. Gorard

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

CE 080328

BEST COPY AVAILABLE

CE

PATTERNS OF PARTICIPATION IN ADULT EDUCATION AND TRAINING

A Cardiff University ESRC-funded Learning Society Project

WORKING PAPER 1

THE REGION OF STUDY

Stephen Gorard
1997

SCHOOL OF EDUCATION

Cardiff
University
of Wales

LIST OF CONTENTS **PAGE NUMBER**

Introduction	1
1. The importance of place and time	2
Wales	3
Industrial South Wales	5
2. Selecting the research sites	10
Nine wards representing the three sites	13
3. The Research Sites	15
Mid Glamorgan	15
Bridgend	17
The three electoral districts	19
West Glamorgan	20
Neath Port Talbot	21
The three electoral districts	23
Gwent	24
Blaenau Gwent	25
The three electoral districts	27
Acknowledgements	27
References	28
Appendices	31

ISBN 1 872330 02 9

PATTERNS OF PARTICIPATION IN ADULT EDUCATION AND TRAINING

A Cardiff University ESRC-funded Learning Society Project

WORKING PAPER 1

THE REGION OF STUDY

Stephen Gorard, School of Education, University of Wales, Cardiff

Grant Holders

Professor Gareth Rees, University of Wales, Cardiff

Professor John Furlong, University of Bristol

Professor Ralph Fevre, University of Wales, Cardiff

INTRODUCTION

This project, funded by the ESRC as part of the Learning Society Programme, is a regional study in industrial South Wales of the determinants of participation and non-participation in post-compulsory education and training, with special reference to processes of change in the patterns of these determinants over time and to variations between geographical areas. The study combines contextual analysis of secondary data about education and training providers with a regional study of several generations of families in South Wales via survey, semi-structured interviews and taped oral histories. The background to the study is further described in Gorard et al. (1996). This paper describes the region of study and provides secondary data used both in the creation of the sample and as a background for the analysis of research findings. The study took place in three markedly different centres in industrial South Wales during 1996 and 1997.

1. THE IMPORTANCE OF PLACE AND TIME

Unlike some studies of participation in employment (Gershuny and Marsh 1994) and previous studies of participation in education and training (Maguire et al. 1993), the region of this study is limited. This is so for two main reasons - the length of time under scrutiny and the need to uncover local determinants of participation. The study encompasses approximately one hundred years of local training and education from the earliest material in the South Wales Coalfield Archive to the present day, and so to some extent the breadth of the study is curtailed by the depth necessary. Policy changes, such as those that affected the Youth Training Scheme, mean that patterns of participation, while predictable to some extent from demographic data and qualifications, and grounded in place, are also susceptible to changes over time (Banks et al. 1992). The patterns and the changes in those patterns which are observed over the hundred years will partly reflect local characteristics. It is the attempt to make these characteristics explicit which also necessitates a very localised study. Education and training opportunities are dependent on local labour markets as much as larger economic conditions (Ashton et al. 1990) and those local opportunities affect the choice of educational pathway from ages 16-19, at least in terms of work-related training (Roberts and Parsell 1988). Since social choices and behaviour tend to be inertial for a number of reasons (Dale and Davies 1994) it can be assumed that these early and locally determined "launchpad" pathways will have a longer term effect on the patterns of participation in the region of study. Local determinants, although changing over time, will in general continue to play a part in the shaping of each individuals learning trajectory, mostly reinforcing their inertial tendencies and so affecting the outcomes of training (Dolton et al 1994) but occasionally producing a shift through sudden radical change. Social changes, patterns of provision, and even the form of implementation of wider policies have local variations and so the importance of place needs to be emphasised in the study and it is only by limiting the study geographically that the necessary level of historical and background detail can be established. However, having identified the patterning in a local context it should then be possible to generalise the findings to a national context more safely than with many wider ranging samples since the merely local effects are now more clearly visible and can be removed from the analysis if desired.

Although some studies have suggested that location and the local labour market are negligible factors in the determination of young people's educational decisions and transitions to work, these are usually because they deal with a select group of people, or a more detailed level of analysis than is usual. For example, Roker (1991) found that many students at private schools had careers planned in national organisations, and so the local labour market played no part in their decision to stay on at school after the age of 16, while Tremlett et al. (1995) suggested that the differences between urban and rural areas in terms of the take up of later learning were more demographic than geographic. On the other hand many studies of employment and training acknowledge the importance of location, even if their research design does not readily allow them to isolate it as a confounding variable in normative analysis (Ashton and Maguire 1986, Roberts 1987). A national survey gives an overall picture, but each locality has its own education and labour market and distinct economic and political culture, which mediate the influence of national systems and policies (Banks et al. 1992). The 16-19 project took place in four areas, and major differences were found between them in terms of the frequency of the five model post-school trajectories of the participants (Bynner 1989), particularly in the proportion of those employed and those undertaking Youth Training in each area (Banks et al. 1992). The chance of someone moving from a period of YT to another job was dependant to a large extent on the area (Roberts and Parsell 1989). The local labour market is particularly relevant for males, but within one locality, such as a city or Travel-To-Work-Area, no difference is discernible (Banks et al. 1992). It is therefore parsimonious to assume that local markets for jobs, and by implication training, are totalities in themselves, providing a coherent unit for analysis.

Wales

Historically there are several gaps in the official statistics for Wales, because although they may have been collected they were never analysed at that level, since no-one needed to make decisions or answer questions specifically about Wales (Williams 1983). In fact, because of the remoteness of central Wales, South Wales has probably been more closely associated economically and culturally with the west of England than with North Wales. Therefore this brief account cannot be definitive and the figures quoted are only intended to be indicative of overall trends.

The average population density in Wales is only 1.4 persons per hectare, compared to 2.4 for Britain as a whole (OPCS 1993). Much of Wales is still predominantly rural in nature, particularly Gwynedd, Powys, and Dyfed (apart from the Llanelli area), where the density is as low as 0.2 persons per hectare. Only 12 towns in these three ex-counties have a population exceeding 5,000, and transport facilities are generally poor (Aberystwyth Policy Group 1990). Motorways, dual carriageways and Inter-City train services are generally confined to the north and south coastal regions and local rail routes such as the Heart of Wales line, have a diminishing service, being more suitable to tourists or shoppers than commuters. Despite the isolated nature of parts of Wales, the population itself is in a state of flux. There was population decline in Wales in the 1970s to 1980s chiefly due to out migration (Istance and Rees 1995). The 1991 census revealed that only 77% of residents were born in Wales, and that this population was not reproducing itself (OPCS 1993), although the overall numbers are now rising chiefly due to greater in-migration (Aberystwyth Policy Group 1990). There is a steady outflow of young people from rural areas, and an equivalent inflow, the majority of whom are elderly and retired English people (Aberystwyth Policy Group 1990). The population is therefore ageing, but not particularly so compared to other parts of the UK. One third of the population was aged 50 or more in 1992 and this proportion is predicted to shrink to 37% by 2006 (Istance and Rees 1995). Thus, it is the middle-aged cohort which is growing in size. Despite the fact that economic regeneration in the region, the arrival of Japanese factories, and EU support has meant an increase in non-Welsh in-migrants (Webster 1990), ethnic diversity is not as obvious as in other parts of Britain. In Wales, 98.5% of the Welsh population are defined as "white" (OPCS 1993), compared to an average of 94.5% in Great Britain as a whole.

Wales had a population of 2,913,000 in 1994, of whom 49% were male, 21% were under the age of 16, 59% between 16 and retirement and 20% retired (Welsh Office 1995$_b$). The civilian workforce is 1,279,000, of whom 965,000 are employed. Roughly 90% of these jobs are permanent, and this figure is fairly even across Wales and across sectors (Welsh Office 1995$_c$). In addition 181,000 are self-employed, 16,000 on a work-related government programme, and there are 117,000 claimant unemployed (see Appendix A). A sharp drop in the number of people employed in mining and quarrying since 1987 has been accompanied by a small rise in production, construction, services, and manufacturing. Male

unemployment is 12.6% and female 4.8%, with both figures decreasing since 1984. The personal disposable income of people in Wales, and consumer expenditure are rising, both in real terms and compared to UK average, but both are still only around 90% of the UK average (Welsh Office 1995b). In many ways the Census 1991 reveals similar levels of disadvantage in Wales as in the rest of Britain (OPCS 1993). The rate of unemployment and number of lone parents are close to the national average and while homes in Wales are less likely to have central heating and private bathrooms they are more likely to have a car than those in the rest of the UK.

However, some observers have reported that comparative social disadvantage in Wales is growing (Istance and Rees 1995). The number of unemployed may be officially underestimated, since 16/17 year olds are often hidden as "status zero", while according to the TUC around 50,000 men in Wales of working age may be economically inactive and not registered as jobless. Of the 459 local authority districts in Britain, 9 of the 10 with the lowest activity rates for men, and 7 of the 10 for women, are in Wales. Deprivation appears particularly in the old working-class towns, mining communities, new suburban housing estates, and agricultural areas (Istance and Rees 1995). In 1993, unemployment in Blaenau Gwent was 14.2%, in Merthyr it was 16.4% and in Aberdare 20.1%.

In addition, even for those in work, the weekly earnings in Wales are the lowest in the UK (Istance and Rees 1995), whereas in the past they have been among the highest, perhaps due to the nationalisation and unionisation of major local industries. Such figures may be worsened by a trend towards insecure employment, particularly effecting the most vulnerable in Wales who might include those living in communities suffering as a result of recent industrial restructuring, ethnic minorities, those with disabilities, those with low skills and single parents (Istance and Rees 1995). There is inadequate child care in a region where 17% of families are single parents. Pay of women lags behind that of men and there are perhaps proportionately fewer women in senior positions in Wales than in the rest of the UK.

Industrial South Wales

A combination of practical and theoretical reasons can be used to justify the use of South Wales in particular as the focus for the study. In principle, once the

importance of a localised study is accepted, anywhere is as good as anywhere else. Each region may display a different pattern and have an interesting story to tell. Wales, which is often bracketed with England in both educational policy and research is in fact markedly different to its near neighbour in terms of education and training in several respects (Istance and Rees 1994, Gorard 1996$_a$). It has also previously been left out of large, otherwise national surveys such as SCELI (1991). South Wales is an ideal setting in many respects since it has witnessed remarkable changes over the last hundred years in industrial and employment structures, forms of work, family relations and wider cultural relationships. The miners' libraries and workers' institutes, and the archives of the South Wales Coalfield Archive suggest that the development of a learning society, at least for men, has been distinctly non-linear with the majority of the work-force in previous generations having opportunities and training provision no longer available or appropriate in South Wales today. Economically, South Wales has changed from being a thriving and expanding area in economic and business terms, to one in which the majority of the working population worked in nationalised industries, such as railways, coal, and steel, to one in which the levels of deprivation and economic activity are relatively high. Recently however, there have been suggestions that one future for South Wales is to become a European mini-"tiger" economy. It is likely that a study of education and training during these relatively condensed periods of dramatic economic and social change will find a link with local conditions more readily than in more stable areas such as South East England. In addition, the local knowledge and contacts of the researchers, and the ready involvement of the relevant Training and Enterprise Councils make South Wales ideal as the choice of a social laboratory.

Industrial South Wales is an officially recognised area measuring 80 miles from East to West, and 30 miles from North to South (Humphrys 1972). It contains 20% of the total area of Wales but contained 70% of the population in 1970. Industrial South Wales was defined by this study as the major parts of the old counties of Gwent, Mid Glamorgan and West Glamorgan, equivalent now to the unitary authorities of Swansea, Neath and Port Talbot, Bridgend, Rhondda Cynon Taff, Merthyr Tydfil, Blaenau Gwent, Caerphilly, Torfaen, and Newport (see Appendix L, Map 1). This area excludes Cardiff and its immediate environs, which although an industrialised port, is a special case and protected to some extent from changes in economic climate. As the capital of Wales, it contains

administrative departments and services on a scale not encountered elsewhere. The seaside towns benefiting from tourism, and other parts of the Vale of Glamorgan combining small-scale farming with residence for commuters travelling to work in other areas are also excluded. Powys and Dyfed are excluded from the focus area, since they are generally sparsely populated with a history of agriculture and some deep mining, rather than heavier industry, and although they have a part to play in defining patterns of training in the region, resource constraints encourage a focus on the more densely populated regions around the M4 corridor.

Industrial South Wales has a regional coherence linked by a history of heavy industry, of coal-mining, iron and steel production using that coal, transport to the ports of the south coast, and the associated smaller industries and services. Overall, industry is now in decline and the entire focus region has been designated an Assisted Area. It is also geographically linked by the coast, by the parallel valleys running northwards at right angles to the coast, and by the motorway and InterCity routes parallel to the coast, while the region is separated to the North by the Brecon Beacons, and to the east by the Severn Estuary and the border with England. However, the focus area is far from homogeneous, containing local specificities which may have significant effects on the patterns of training encountered. It is therefore necessary to characterise the local contexts and select a combination of study sites that accurately reflect the diversity (see Appendix B). It is not possible to obtain a random, or stratified sample of the whole region since the project requires the identification of local determinants of participation in education and training, as these will help to explain any local patterns observed, which in turn requires detailed local knowledge. Concentrating on as few sites as possible allows a greater depth of knowledge about each. Having selected local sites to stand for the whole region, and identified specific local factors, these factors can then be isolated and a more general picture should emerge. Thus, ironically, by concentrating on local processes and their outcomes it may be possible to identify more general social processes taking place in South Wales of relevance to other areas displaying different patterns.

It is important not to overestimate the length of time that industrialisation has been in South Wales (e.g. the Rhondda coalmines functioned for less than a hundred years). Many workers were in-migrants from rural areas, and so the

area remains very much influenced by an agricultural ethos, while the majority of families, such as those in the valleys, who moved here to provide a workforce for the "new" industries are themselves relatively new to the area. Coal, iron ore, limestone, silica and water were the bases of industrialisation in South Wales which rapidly became a boom area in the nineteenth century (Humphrys 1972). Although towns like Bridgend and Neath predate this period, and even today areas like the Vale of Glamorgan, Gower and Monmouth are chiefly agricultural, the pattern of settlement owes more to industry than anything else (Humphrys 1972). Until 1918 the local immigration rate was similar to that of the USA. The population in the Rhondda valleys was 4,000 in 1861, but thirty years later had grown to 163,000 due to the expansion of the mines, which produced instant communities of ribbon-like developments (Nash et al. 1995). The coal valleys comprise individual villages clustered around a mine, and the shortage of building space forced the developments to be along the floor of the valleys, and only a short distance up the sides (Keen 1979). Tinplate production was a local boom industry due to the growing prevalence of food canning in the USA starting in the 1870s, and South Wales had a virtual monopoly in this area. However, this monopoly was broken by the discovery of tin in the USA in the 1890s and the subsequent protectionist policies (Howell 1988). Many of the affected factories therefore converted with some relief to armaments production in 1914. The dominant industry was coal, with 56.8 million tons of coal produced in 1913. In good times the wages in mining were good and there was a high demand for men. The influx of new workers deluged many traditional aspects of the region, and mining communities developed their own distinctive life styles. South Wales had the fastest growth rate in the UK, with in-migrants often from other parts of Wales moving to Gwent and Glamorgan. Then came the Irish and English, leading to a net gain of 129,000 from 1901-11 and the consequent destruction of some Welsh-speaking areas (Howell 1988), especially those to the east.

In the 1920s, as the first US companies began investing in Wales (Wales "2000" 1996), steel manufacturing away from the coast at places like Blaenafon, Tredegar and Ebbw Vale collapsed due to transport costs and competition, along with the continued decline of tinplate. Even the coal industry lost many of its markets to US rivals during the Great War and the blockade of the UK (Howell 1988). Whereas there were 234,000 miners in 1913, there were only 136,000 by 1939. Into all of this came the General Strike of 1926. By the 1930s

unemployment in Tredegar was 27.4%, in Newport 34.7%, and by 1935 over 50% in Merthyr Vale (Howell 1988). Consequently between 1921 and 1939 400,000 people left South Wales, causing the age profile of those remaining to rise as the leavers were mostly young people (Humphrys 1972). Unemployment disappeared temporarily during the wars of 1939-45, but the underlying industrial decline continued.

Between 1945 and 1970, the pattern of industrial life in South Wales was transformed. By 1970 nearly everyone was in a job that did not exist in the same place in 1945 (Humphrys 1972). Many companies at that time were large employers, with little diversity in industrial classification. The mines were nationalised in 1947, while industry became more and more concentrated on the southern coastal strip (Howell 1988) and the numbers employed in mining dropped again from 112,000 in 1944, to 106,000 in 1960 and only 60,000 by 1970. In the main, local companies were not owned by Welsh people, who had only a small traditionally established middle-class. The major decline of the coalfield took place between 1959 and 1970 when 86 collieries closed (Rees and Rees 1983). Some of the miners moved elsewhere, and some of the younger ones moved to new industries in the south or the mouths of the valleys. Many of the old industrial towns became little more than dormitory centres, but due to inertia few people moved to where the new jobs were (Humphry 1972). The New Towns Act 1946 enabled the creation of new towns like Cwmbran in Gwent in 1949, to replace the existing poor housing.

By 1971 60,000 people in South Wales worked for US companies, such as Hoover, and by 1979 there was so little industrial disruption in the area coupled with low wage and land costs and a package of inducements that still more overseas companies invested in the area, including those from Japan and Korea, although the US companies are still the most numerous (Wales 2000, 1996). These used South Wales as a springboard from which to supply the rest of the continent. Mining had so far declined that only 30,000 were employed by 1979 (Howell 1988). Partly in response to this, British Coal Enterprise was founded in 1985 to train ex-miners for job and career changes (Rees and Thomas 1991). In the 1980s the numbers employed in coal mining continued to fall, but there was an increase in self-employment and small businesses in South Wales, coupled with an increase in women's employment. However, gender roles have generally been

highly segregated in mining families with women confined to servicing their mining husbands and sons (Rees and Thomas 1991). Miners have consequently been seen as archetypal proletarians and male chauvinists who are reluctant to tolerate their wives as breadwinners, or to lose benefit because of their wives' employment (Hebron and Wykes 1991).

The population in South Wales today is not reproducing itself. Younger and more qualified people are leaving the area so that the population is both ageing and declining in economic activity. However the population of towns like Bridgend is growing, chiefly through in-migration. The rate of economic activity for men is declining and that for women is growing, in line with national trends, but the decline for men is more pronounced and the increase for women less so. Again, towns like Bridgend with relatively high levels of activity for women, mostly in part-time work, are different to the mining valleys where the activity rate for women is and always has been very low. Very few people are self-employed and even fewer are on government schemes, although there are more in Port Talbot. Unemployment is higher than for the UK, and the occupational class and education profiles are generally lower. The clear differences between the sites and regions reinforce the importance of using location as a context for analysis.

2. SELECTING THE RESEARCH SITES

Since the study conceives patterns of participation in education and training as socio-economic trajectories over time (Gorard et al. 1996), and uses primary data from at least as far back as 1933, both the current and the past characteristics of the study region need to be taken into account in selecting the representative study sites. While in some senses this makes selection of the sites for the study more complex, it is also true that any sampling strategy cannot be valid for all of those 50 years, and so if the stratification is less than perfect for current indicators it may not matter as much as for a typical cross-sectional survey. What is needed as much as a statistical snapshot of the potential sites at any moment in time is a sense of how learning-related characteristics, such as employment and qualifications, have changed in the past, and what the current trends are. Also, since the study is of local determinants of participation in training, all other factors being equal, preference will be given to those areas that have had a relatively stable population in terms of migration over the period of the study. In general, other than those areas identified above - sparsely populated with a

history of agriculture, Cardiff, and the seaside towns - the region can be seen as comprising three distinct local types. The valleys, particularly those to the east and centre of the area have a history of coal-mining with allied manufacturing from the nineteenth century onwards. They are heavily urbanised and densely populated in pockets, although farming continues on the hills between them. These sites are typified by Aberdare, Rhondda, Merthyr, Rhymney and Blaenau Gwent, and most are in industrial decline (Welsh Office 1996). Heavy manufacturing industries developed after 1945 in some sites along the coast. These were urban areas, primarily producing petro-chemicals, metals such as steel, and steel-based products. These industries, like the coal-mines, are also declining but the decline has not been as dramatic as the large-scale closure of the pits. Such sites appear in Llanelli, Swansea, Port Talbot, Neath, and Newport (Welsh Office 1996), although Newport like Cardiff has some new development giving it a hybrid character. New manufacturing industry tends to be closer to the M4, and so slightly further from the coast. Relatively new factories, many built by overseas companies, are causing some urban areas to expand, and the communities that are springing up as a result suggest one possible industrial future for South Wales - that of providing a relatively skilled but relatively cheap labour force for foreign investors. These towns are typified by Bridgend, Pontyclun, Llantrisant, and Newport, and some of the settlements surrounding the Heads of the Valleys road (Welsh Office 1996).

Thus, three sites can reasonably be used to typify the focus area: one in the heart of the depressed coalfield, one in a large urban centre historically linked with heavy industry, and a third in a rapidly expanding town south of the coalfield, of which Bridgend is the clearest example. Bridgend also allows examination of the recent impact of Japanese investment in the area, which may be part of the reason for the continuous rise in population since 1981 in the area covered by the new unitary authority (Welsh Office 1996). The balance of new and closed manufacturing since 1994 in Bridgend is good, unemployment in 1995 is markedly less than in 1993, and the basic standard of education as assessed by GCSE benchmark (37%) and failure rate (10%) is good for the region. More than half of the work-force is female (56%), unlike in most other areas of South Wales, and a large proportion is employed part-time (32%). Newport shows a similar picture in some ways - continuous population increase since 1981, net gain in manufacturing 1994-1996, and a decrease in unemployment 1993-1995

(Welsh Office 1996). However neither the employment rates nor the basic educational attainments at age 16 (35% benchmark, 14% failure) are as high as Bridgend, and Newport has markedly fewer part-time (24%) and female workers (46%) and so is more like the other areas of declining or stable manufacturing, such as Port Talbot.

The selection of Bridgend, which is in the area dealt with by the Mid Glamorgan Training and Enterprise Council effectively determines the other two study sites. Since Bridgend is in Mid Glamorgan and can therefore be taken to represent the area covered by Mid Glamorgan TEC, and since Gwent and Mid Glamorgan contain the archetypal coal valleys, Blaenau Gwent was chosen to represent the depressed coalfield. Blaenau has had a continuously falling population since 1981, in complete contrast to Bridgend, with a slight decrease in unemployment recently, a still predominantly male workforce (53%), and poor raw-score educational measures (28% benchmark, 14% failure, Welsh Office 1996). However, there may be an indication of a very recent slight upturn in population, and the balance between new and closing manufacturing industry has been maintained from 1994 to 1996. In these respects the picture in Blaenau Gwent is very similar to that in the unitary authorities in what used to be the mining valley areas of Mid-Glamorgan, such as Rhondda, Cynon Taff, Merthyr Tydfil, and even Torfaen (Welsh Office 1996). All show a slight drop in unemployment 1993-95, all are holding their own, or even showing a slight net gain in manufacturing 1994-96, with about 50% of the workforce female and about 25% part-time. Thus Blaenau, although in what used to be Gwent can be taken to represent the coalfield. It also has the distinction of a very poorly qualified population, worse even than Merthyr Tydfil (Welsh Office 1996).

Thus, the manufacturing site was in the West Wales TEC area, and although this left a choice of Llanelli, Swansea or Neath Port Talbot, the latter was preferred for its proximity and the detailed local knowledge of the members of the research team. All three areas have shown minor fluctuations in population since 1981 (Welsh Office 1996), but whereas Swansea gives a picture of some new economic growth similar to Bridgend - falling unemployment 1993-95, high basic educational standards (46% benchmark, 8% failure), and high proportions of part-time (30%) and female employees (53%) - Swansea has shown no gain in population or manufacturing 1994-96 (Welsh Office 1996). Neath Port Talbot,

on the other hand lies somewhere between Bridgend and Blaenau Gwent on several measures - slight drop in population 1981-91, slight overall loss in manufacturing 1994-96, slight drop in unemployment 1993-95 and median basic educational qualifications (33% benchmark, 12% failure). Thus although it still has a predominantly full-time (78%) and male workforce (58%), Neath Port Talbot is in marked contrast to both other sites.

The next stage of the sampling process was to define the contours of the three regions of study. Since most of the data used for initial selection was at the level of unitary authority and the electoral register was to be used to select the households, the three sites were defined in terms of electoral divisions (wards) within each authority. In the new unitary Authority of Blaenau Gwent, all of the electoral divisions are in the coalfield area, and all sixteen were considered as candidates (see Map 2, Appendix L). In Neath and Port Talbot unitary Authority, the wards to the north and east of the main conurbations are agricultural, or sparsely populated areas which are not typical of industrial South Wales, or coalfields which are already represented in the sample by Blaenau Gwent. None of these were considered as candidates. Neath and Port Talbot, for the purposes of this study was therefore defined as a coastal strip of eighteen wards (Map 3). In Bridgend Unitary Authority the wards to the far west of Bridgend town and the wards to the north were also excluded by the definition of the study region, since they are sparsely populated and agricultural, or seaside towns. Bridgend is here defined as a conurbation of thirteen wards (Map 4).

Nine wards representing the three sites

The same argument about restricting the focus area in order to obtain knowledge in depth is also used to justify selecting only a few wards in each site. The nature of the sample is described more fully in a later paper and only a brief outline of the strategy is given here. With a total sample size of around 1000, the sample for each site is only 333 whereas each site has a population of some 80,000, so it would be unreasonable to expect a random sample to be perfectly representative of the expected diversity. The strategy employed was to select three wards in each site which were stratified in terms of the educational level of the population, and which encapsulated the variety of communities encountered. However there were other practical reasons for dealing with the ward level. If the whole of the site had been randomly or systematically sampled the cost of the electoral registers

for all wards, the cost of the travel for the many calls to each household that the survey interviewer needed to make, and the cost of the subsequent in-depth interviews would have been much larger. The wards are not intended to be a unit of analysis at any stage, since the numbers in each would be too small. They were selected partly on the basis of secondary statistics and partly for qualitative reasons such as the mixture of housing, and the type of local employment. One problem for the sampling was that although the proposal talked in terms of counties, and data on population characteristics is available for TEC areas and for counties and electoral wards, local government has been altered since 1/4/96 under the Electoral Arrangement Orders 1994. In many cases this has simply meant the allocation of existing electoral wards to different unitary authorities, but in others the new authorities do not follow the borders of previous wards.

Since the study is primarily concerned with education and training, it was appropriate to use educational data to select wards to be representative of the level of education at each site. The only reliable data on education available at this level of disaggregation was the ward level 10% sample linked to the 1991 Census, and this only showed the proportion of inhabitants with a qualification from higher education, such as a diploma or degree. This is a useful measure since it affects the possibility of further continuous education, and is clearly linked to both occupational class and economic activity rates. However, the fact that it is a different measure to the GCSE benchmarks used in selecting the sites, and that it applies to only a minority of the inhabitants means that it is only used as a starting-off point, in an attempt to encapsulate the range of educational experiences and attainment in the study region. Candidate wards in each site were selected representing those with a high proportion of HE qualified inhabitants, those with a low proportion and those in the middle (see graphs in Appendix M). These candidates were then compared in terms of other factors such as size, migrancy, and the economic activity of the population, and visited by the researcher to obtain a clearer understanding of their character. All other things being equal the wards were chosen to be as even in size as possible, using the number of people usually resident at the time of the Census (1991), both within and between research sites so that the sampling fractions were as similar as possible, and to reduce the need for expansion weights (Lee et al. 1989). As some of the data from the survey is analysed as a retrospective longitudinal study, the population of each ward was as stable as possible, using the proportion of

residents having a different address one year before the Census (1991) as an indicator. In this way the survey is unlikely to be swamped by newcomers to South Wales who although interesting in their own right, cannot really elucidate local parameters in the formation of learning identities and trajectories. Since one focus of the study is on work-based training, and another is on those in the workforce excluded from training, and the sample is to be taken of adults before the age of retirement for men, it also made sense to consider the proportion of the population who were economically active at the time of the Census (1991).

In summary, three wards were chosen in each site, each with differing levels of educational qualification, with as many households as possible, as many still economically active people as possible, and as few migrants as possible, of similar size within each site, with total populations of similar size for each site, and with the overall sample as close to the characteristics of industrial South Wales as possible. Since in some cases these principles were contradictory, several solutions were possible. The nine wards described at the end of this paper were perhaps the best possible compromise. All figures, unless otherwise referenced are percentages calculated from data in the 1981 and 1991 Censuses provided by the National On-Line Manpower Information System in Durham.

3. THE RESEARCH SITES

Mid Glamorgan

The first research site, Bridgend, is in the old county of Mid Glamorgan which is one of the poorest parts of South Wales with perhaps the worst record in terms of educational attainment. Its residents currently have a very low economic activity rate with a high proportion permanently inactive through illness (Mid Glamorgan TEC 1996). Despite a relatively high manufacturing base, with proportionately more employees working in Manufacturing than anywhere in Wales, and more than average for the UK as a whole, it has a low average Gross Domestic Product, and this figure has actually decreased in every year from 1989 to 1994. The suggestion that this high figure for manufacturing is not due to industrial strength but because there are few other jobs in the area is reinforced by the lack of managers and employees of administrative grade. The most common occupation is as a plant or machine operative (Mid Glamorgan TEC 1996). There is a large amount of youth unemployment, and more than 20% of the unemployed

have been so for more than two years. In fact, male full-time employment has been decreasing every year since 1981. Why this is so may perhaps be glimpsed in the figures for jobs in mining, which have fallen from 18,700 in 1981 to 1200 in 1993, and for jobs in electricity, gas and water which have dropped from 2400 to 600 (Mid Glamorgan TEC 1996).

Mid Glamorgan also has a low qualification profile, even lower than for Wales as a whole, as evidenced by the proportion of the workforce with each level of qualification, degrees especially. It has the worst school leaver benchmark figure, 30% with 5+ GCSEs (Welsh Office 1995a), and the most school leavers with no qualifications at all (Reynolds 1990). This relatively poor performance could be partly due to the class backgrounds of the school intake, since Mid Glamorgan has the highest ratio of working to middle-class families in Wales (Reynolds 1990). However, this cannot explain all, since South Glamorgan which has the lowest ratio of working to middle-class families, and the most private schools, has the second highest number of school leavers with no qualifications. For both Training For Work and Youth Training, Mid Glamorgan has one of the least successful records in Wales in terms of completion and qualification rates, and is in these respects significantly below the average for Great Britain (DfEe 1995). It is therefore perhaps not surprising that there are reports of shortages of suitable applicants for some positions, especially for skilled manual tasks such as machine setters, fitters and toolmakers (Mid Glamorgan TEC 1996). These local skill shortages may be more acute as a growing number of residents travel to work each day in neighbouring districts such as South Glamorgan, the relatively affluent neighbour of Mid Glamorgan. In the 1996 Annual Labour Market Assessment, many businesses had vacancies at the time of the survey, but that this is evidence of skills shortages rather than boom is backed up by the finding that 40% of large companies reported difficulties in filling their vacancies despite the relatively high local unemployment level. Only 13.2% of local employees had received any job-related training in the four weeks prior to a recent survey, and this was lower than the UK average and much lower than the 17.5% reported by its neighbour South Glamorgan (DfEe 1995).

However, this relatively depressing situation should not be exaggerated. The average GDP is projected to increase in the years 1994-99. Overall, the level of

unemployment is falling slightly, for example from 23,137 in 1994 to 20,756 in 1996 (Mid Glamorgan TEC 1996). Since nearly two thirds of the current unemployed are aged under 35, it is not surprising that the numbers staying on in post-compulsory education are increasing. Whereas 49% of the then Year 11 pupils stayed on in full-time education after the age of 16 in 1990, by 1994 the figure had risen to 66.1%. However, the selectivity of this change is seen in the fact that the increase in numbers in education has led to a reduction of those in employment and training, but not those in unemployment (Mid Glamorgan TEC 1996). The youth unemployed are still leaving school at 16, and it is those who in previous generations might have left school for a job or Youth Training who carry on to sixth form. The level of qualifications among the residents are also slowly rising. More 16 year olds are now staying on in full-time education in Mid Glamorgan (63.3%) than in the UK as a whole (60.9%, DfEe 1995)

There has also already been growth in a few occupational areas. Since 1981, occupations with classifications such as Hotels, Public Administration and Health have shown a slight increase in the region. Also, despite the overall decrease in male full time employment, part-time employment in the region, especially for women has increased (Mid Glamorgan TEC 1996). This pattern is similar to that observed in the whole of Wales, but is more clearly marked in Mid Glamorgan than anywhere else. Why this might be so is partly explained by one of the regions of Mid Glamorgan that have bucked the trend - Bridgend.

Bridgend

Bridgend is a town and the name of the new Unitary Authority centred on the town. It is smaller than the travel-to-work area which includes Bridgend town and the Llynfi, Garw and Ogwr valleys including Maesteg to the north, plus the coastal plain with seaside areas such as Porthcawl. Originally a market and agricultural town, Bridgend grew to service the iron and coal industries to the north in the nineteenth century, becoming a focal point for transport networks (Boddy et al. 1990). Work began on an ordnance factory here in 1938, which employed 34,000 people, mostly women at its peak during the war of 1939-45. Much of this factory has now been converted to Bridgend Industrial Estate, where over 75 firms operated by 1950, employing over 3,000 people. By 1967 this estate had extended to 200 acres. At the same time, in the north part of the current travel-to-work-area the coal industry declined as rapidly as it had grown.

In the mid-1920s there were 409 local collieries employing 40,000 men but by the 1970s there were only 7 working collieries employing 6,000 (Boddy et al. 1990).

Its population, unlike the two other research sites, has increased from 1981-96 partly due to in-migration but mostly because of a relatively high birth rate coupled with a low death rate, the highest and lowest respectively in South Wales (Welsh Office 1996). Bridgend is very unusual in this respect. However, the live birth rate is still falling in common with the rest of Wales. The Welsh Office also suggests that Bridgend shows a slight decrease in unemployment which is reasonably common, coupled with a low rate of long-term unemployment which is not. Its workforce has a relatively high proportion of female and part-time employees. It is the site of several multinational companies and overseas investments, including Japanese firms such as Sony, and north American firms such as Acco UK Ltd, Align-Rite Ltd, Biomet Ltd, Carters (J and A) Ltd, Ford Motor Co. Ltd, Harman-Motive Ltd, and Jamont UK Ltd ("Wales 2000" 1996). In fact, 5,200 people are employed in 17 overseas plants in the study area (Welsh Office 1996).

The proportion of residents with qualifications beyond A level is relatively high for South Wales and rose from 7% in 1981 to 9% in 1991, with an average 14 A level points per candidate in 1995 (Welsh Office 1996). In 1995, 37% of the relevant year groups gained five or more GCSE grade C, but a relatively large 11% did not gain even one CoE. The class profile of the area is also changing, with skilled manual jobs being replaced by associate professional ones. Only 5.4% of the population reported being able to speak, read or write Welsh, perhaps because in one year alone 8% of the inhabitants has moved home, a quarter of them from outside the area.

However this picture of growth must not be exaggerated. No clear drop in unemployment is visible between the snapshot dates of the two last Censuses in 1981 and 1991, but the economic activity rate for men has significantly decreased. New jobs have tended to be filled by women (Boddy et al. 1990), but while the activity rate for women has increased, the two changes are not commensurate. While full-time male employment in Bridgend Travel-to-Work Area dropped by 16% from 1981 to 1987, female full-time employment rose by

8%. Female part-time employment rose 27% over the same period. There has been an increase in commuting into Bridgend, and Port Talbot, from the areas to the north (Boddy et al. 1990). Despite the increase in population and decrease in unemployment, there may not actually be any more jobs in Bridgend than there were 15 years ago. The pattern for employment is remarkably similar in all three sites. A clear majority of those in work are employed full-time, with part-time work much more common for women. Very few people, especially women, claim to be self-employed, while even fewer are on government schemes.

The three electoral districts

The expanding town of Bridgend is represented in the fieldwork of this study by three electoral districts which have been chosen to represent the range of social, especially educational, conditions of the site. The populations of the three districts are similar, between 5,224 and 5,755 in 1991. In general, the district with the most highly qualified workforce also has the most stable population but a larger number of residents who have moved from outside the region, a larger proportion of residents from recent ethnic minority backgrounds. It has larger houses, less industry on site and the residents are in more professional occupations. The opposite conclusions can be drawn about the districts with the least qualified workforce, and these conclusions hold for all three research sites.

Coity Higher is a suburban district in south central Bridgend, with some large semi-detached houses and a few shops. A clear majority of the residents are professionals or associate professionals and 34% of the workforce have a higher educational qualification such as degree or diploma. Coity is also the only district in which a substantial proportion of inhabitants moved there from outside the UK in the year preceding the census.

The Newcastle district is based upon a housing estate and has a mixture of modern and council owned property and older semi-detached houses. It is in central and west-central Bridgend. The most common occupational classification of residents is associate professional but there are also a large number of skilled manual workers and nearly 28% are educated to degree level or equivalent. It has the lowest economic activity rate for both men and women of the three districts in the study.

St Bride's Minor is in the north, on the outskirts of Bridgend town. It is therefore much less densely populated. The housing is a mixture of new developments and semi-rural. It has a high rate of economic activity, but also a high rate of male unemployment and only 11% of the workforce are educated to above A level standard.

West Glamorgan

The second research site, Neath Port Talbot, is in the old county of West Glamorgan. It is part of the Annual Labour Market Assessment for West Wales, which as well as Swansea and Llanelli includes relatively remote and sparsely populated parts of Wales which are clearly not part of industrial South Wales. In some respects West Glamorgan has similar problems to Mid Glamorgan, being also within reasonable commuting distance of the relatively affluent South Glamorgan. Investment capital to the region is falling (West Wales TEC 1996). Its residents have a low economic activity rate, and the levels of personal income and average Gross Domestic Product are both significantly lower than the UK. Most local companies are in the Service sector, not manufacturing, and most have fewer than five employees (West Wales TEC 1996).

Around one third of employers reported problems in areas of recruitment, and not just for skilled posts (West Wales TEC 1996). The problems were especially to do with basic skills such as numeracy and literacy, customer care, and unfavourable attitudes to work. Only around 14% of firms had a written training plan and only 13% of employees had received any training in the four weeks prior to the survey, lower than the national average. The proportion of 16 year olds remaining in full-time education is the lowest of the three research areas (56.5%) and much lower than for the UK (60.9%, DfEe 1995). With Mid Glamorgan, West Glamorgan also has the largest proportion of school leavers gaining no qualifications at all (Reynolds 1990). However, West Glamorgan is generally not as economically depressed as Mid Glamorgan. The qualifications of the residents as a whole are around average, or perhaps slightly better than for Wales, perhaps due to in-migration or better adult training opportunities in the area.

Neath Port Talbot

The new Unitary Authority of Neath Port Talbot is centred on two neighbouring conurbations. Since it also includes the Lliw valley and surrounding areas, the overall population density is low, but these areas do not form part of this study and are excluded from further analysis as far as possible. The population is not reproducing itself, dropping in the years 1981 to 1992 (Welsh Office 1996), partly due to out migration from the area, and although the level of long-term unemployment in Neath is reported to be low, the rate of unemployment in Neath Port Talbot as a whole is only decreasing slowly. The economic activity rates are low, even for South Wales. A high proportion of the jobs available are in manufacturing, in fact the highest proportion in Wales with the lowest proportion of service jobs which are otherwise the most common form of employment in West Wales (West Wales TEC 1996). Consequently there is a low proportion of female and part-time employees in Neath Port Talbot (Welsh Office 1996). It is also the authority with the highest proportion of Welsh language speakers, 10.6% able to speak, read or write Welsh. 7% of residents had a different address in 1991 to 1990, of whom less than 2% moved from outside the district.

Neath Port Talbot has fewer overseas investors than the other sites, relying until recently on nationalised industries. The north American firms locally tend to be smaller concerns than those in Bridgend, Borg Warner Automotive Ltd, Texturing Technology Ltd, and TRW Steering Systems Ltd ("Wales 2000" 1996). In all, 3,950 individuals are employed in 15 overseas plants (Welsh Office 1996).

As in Bridgend, but to a greater extent, the class profile is changing with fewer skilled manual jobs and more associate professionals. Although this is clearly an important change for the area, it may be partly caused by a natural "inflation" of job titles taking place between the Censuses. The proportion of residents with qualifications above A level equivalent is low but rose slowly from 5% in 1981 to 6% in 1991, with 13 A level points per candidate. In 1995, 33% of the relevant year group gained five GCSE grade C, while 9% gained no qualifications at all. In the Annual Labour Market Survey a particularly large number of firms in this site reported problems of literacy and numeracy among potential employees (West Wales TEC 1996). It has the lowest proportion of students continuing to full-time education after 16, 67%, in the area, but it does have the largest numbers on Youth Training or training credits (West Wales TEC 1996).

The town of Neath, unsuccessfully proposed as a site for a university in the 1850s (Eaton 1987), has grown and shrunk dramatically this century. Its earliest industrial development was for the reduction of imported copper ores and later for lead. Iron, coal, tinplate, and oil-refining also played a part. The Metal Box Company settled locally in 1935 to use the tinplate, and the British Aluminium Company arrived in 1938. Its population grew from 14,000 in 1901 to 32,000 in 1951, but by 1971 it had fallen again to 29,000 (Jenkins 1974). It continues to fall partly as a result of economic decline and partly as a result of structural changes in the form of clearance of the central town area. It has been the site of Neath Tertiary College since 1982.

The town of Port Talbot, a concentration of industry and population on a narrow strip of coastal plain (Port Talbot 1996) has an even more uneven history. It grew from 1860 onwards due to industries such as tinplate works, the docks and also small tobacco producers. These replaced the more common cattle and sheep farming previously (Jones 1991). This inevitably led to an expansion of local housing services and transport, and a gas works and drainage system completed in 1869. The docks were enlarged in the 1890s. As well as producing new development, Port Talbot became a magnet for industries inland from places like Cwmavon which were rebuilt at the dockside. It also drew the coal trade from Porthcawl, which has since become a seaside resort (Jones 1991). However, some depression was caused by protectionist policies in the USA at the turn of the century, and it is important not to overestimate the size of town at that stage. In 1901 fewer than 100 men worked in iron and steel, and fewer than 800 in tinplate, with perhaps the same again in mining. The population was only 7,552 which is only half that of Neath at the time. A boom came in the Great War when there was an increased demand for munitions, particularly from the 1901 Port Talbot and 1916 Margam Steelworks (Jones 1991). It is perhaps at this stage that the name Port Talbot, originally used to refer only to the docks came to be used for the east bank across the river as well. The decline in demand and the depression after the war, despite having MacDonald, the local MP, as Prime Minister, led to the closing of the Steel Works in 1931 (opening again in 1934) as well as local collieries. There was 43% male unemployment. All building stopped until another war created another surge in demand, which is why although the houses in the centre of town are Victorian and Edwardian, everything after this is

post-1945, for example in Sandfields and Baglan Road. By 1921 the population of the recently established borough combining Port Talbot, Margam, Sandfields and Cwmavon was 40,027. In 1952 the Steel Company of Wales, one of the first integrated steel works in Britain, opened on Margam Moors. By 1965 it employed 18,000 people and offered 30 apprenticeships per year (Fevre 1989), but the local population was still only 52,000 which may be indicative of the draining effect of the depression. The docks ceased to export coal in 1962 and the old docks closed in 1971. A new Tidal Harbour exporting only iron and steel opened in 1970, which has the deepest berthing facility in the Bristol Channel, and is dedicated to the import of raw material for steel works at Port Talbot and Llanwern (Port Talbot 1996). The building of the M4 motorway and the position of Port Talbot on the InterCity line between London and Swansea opened land communications. It briefly became so prosperous that the steelworks became renowned as "Treasure Island" (Jones 1991). BP Chemicals opened a plant there in 1963, and the Steel Company of Wales has been nationalised and then privatised again. Its employees fell from 18,000 to 4,000 and Treasure Island became "Giro City" (Fevre 1989). By 1996 the steel works were still the largest employer in the area, employing 4,500 (Port Talbot 1996).

The three electoral districts

The three electoral districts chosen to encapsulate the range of social conditions in the established manufacturing site of Neath Port Talbot are not as close in population size as in the other two sites. The largest district had 6,975 residents and the smallest had 4,455 in 1991, and as with the other sites, there are slightly more women than men.

Cimla is a relatively leafy suburb of the town of Neath with large semi-detached houses and off-street parking. The population is relatively stable, showing few house moves from 1990 to 1991, with a high rate of economic activity, especially for women. The population is relatively well-educated for this site, with 19% of the workforce educated beyond A level and a high proportion of non-manual and professional occupations. Unemployment is low.

Bryn and Cwmavon are two separate communities in a valley to the north outside Port Talbot. They are effectively large villages on hillsides and the valley floor with a range of styles and age of housing. Cwmavon is based on a tiny community

which was built up in the nineteenth century by the English Copper Company (Jones 1991). However in the late nineteenth century local industry, such as the steel and copper works moved down from Cymafon and was rebuilt at dockside in Port Talbot when it was enlarged to reduce transport costs. Cwmafon is now a residential commuter area for Port Talbot (Port Talbot 1996). 16% of the workforce have higher educational qualifications, most being in skilled or associate professional occupations.

Sandfields West is a large council housing estate on the sand dunes by Aberafan Beach, very close to the industry in Port Talbot. Originally designed for the workforce of the local steel works, it has very few shops and little variety of housing. Developed from prefabricated housing placed there during 1939-45, it is an "instant community" with a slightly forlorn air (Jones 1991). The residents are relatively stable, as assessed by moves between 1990 and 1991. Very few of the residents are of recent minority ethnic origin. They have a low rate of economic activity, and only 5% of the workforce have qualifications beyond A level. Most of those in work have manual or unskilled occupations. There is high male unemployment, and a significant number of the residents are on government schemes.

Gwent

The third research site, Blaenau Gwent is in what was formerly the county of Gwent adjacent to the border with England. It is a region with a tradition of high profile politicians, such as Aneurin Bevan, Michael Foot, and Neil and Glenys Kinnock representing local constituencies. Like Mid Glamorgan, it is an area of contrasts and contradictions. The agricultural areas like Monmouth apart, it shows two overall patterns of decline and retrenchment inland, and relative boom conditions near the coast around Newport.

There is a steady increase of inward investment in Gwent creating nearly 2000 net new jobs in the first eight months of 1995 alone (Gwent TEC 1996). Gwent now receives £1 of every £3 invested in Wales and is host to a quarter of the multinationals firms operating in Wales, and this is before the promise of the Korean industrial conglomerate, the LG Group, in 1996 to create two new businesses in Newport, LG Electronics Inc. and LG Semicon Co. It has been suggested that these businesses will create 6,100 new jobs and bring investment

worth £1.7 billion to the area ("Wales 2000" 1996), as well as having a knock-on multiplier effect in terms of optimism and the provision of services. Occupations in electronic engineering and assembly as well as financial services are on the increase, so that overall unemployment is reducing in the region.

In education, both the retention rate of students at the age of 16 and the qualification profile of the residents are increasing (Gwent TEC 1996), while in terms of both Training For Work and Youth Training, Gwent has one of the most successful records in Wales of completion and qualification rates, and is in these respects comparable to the average for Great Britain (DfEe 1995). The proportion of 16 year olds remaining in full-time education (59.1%) is similar to that for the UK (60.9%, DfEe 1995). The proportion of employees who received job related training in the four weeks preceding a survey was slightly higher than the national average at 13.6% (DfEe 1995).

However, this upbeat assessment cannot disguise that the decline of the mining industry has had a disastrous effect on some of the communities further from the coast. The population of the area is decreasing, with a drop of around 5% overall between the last two Censuses in 1981 and 1991 (see Appendix I). Most of this decline is attributable to out migration, particularly among those aged 20-35, who have moved to look for jobs. The population is therefore also ageing. Despite this, the pattern of long-term unemployment for many remains unchanged. Economic activity rates are low, especially for women and in the valley communities, while the rate is actually declining for men, from 93% economically active in 1971 to 82% in 1991 (Gwent TEC 1996). These men have in the main been replaced by jobs held by women, of which a higher proportion are part-time in line with the national trend. Similarly, while the educational profile of the population is improving, it is clearly below average for Wales, and overall is probably the lowest of the surrounding areas of the three research sites.

Blaenau Gwent

The new Unitary Authority of Blaenau Gwent is the smallest of the three sites based upon two valleys running north of Newport, "where the valleys meet the mountains of Gwent". It is the site of the proposed first new urban village in Wales, at Victoria where the Festival of Wales was held (Blaenau Gwent 1996). It has a low birth rate and high population death rate, one of the lowest and highest

respectively in South Wales (Welsh Office 1996). The population is therefore falling heavily, and the situation is not helped by the out migration of many young people. It is still the most densely populated of the three sites (Welsh Office 1996), but has the fewest roads. Three of the electoral wards in Blaenau - Nantyglo, Tredegar and Cwm - are among the most highly deprived in Gwent and probably in South Wales (Gwent TEC 1996).The area has very low civilian activity rates, especially for women, and long term unemployment is high (Gwent TEC 1996). Most jobs are concentrated in the manufacturing sector (Welsh Office 1996).

Although it does have some overseas investment including Japanese, German and USA, in an industrial base of over 260 companies of whom 75% are new to the area (Blaenau Gwent 1996), this is concentrated in Ebbw Vale town, with all of the local north American companies located there, Aeromotive UK Ltd, ITT Automotive UK Ltd, and Sears Manufacturing Company Europe ("Wales 2000" 1996). The companies are generally in industrial estates and business parks. In all, 2,400 people are employed in 24 overseas plants (Welsh Office 1996). The class profile has risen from 1981 to 1991, just as with the other research sites, but the similarity of the pattern regardless of other economic indicators confirms that this may be an overall societal change or an inflation of the measure rather than a specifically local phenomenon.

Blaenau Gwent has a poor record in terms of educational outcomes at all levels (Welsh Office 1996), and although the proportion of the population qualified beyond A level rose from 4% in 1981 to 5% in 1991, 5% is still one of the lowest in Wales, with an average 12 A level points per candidate. 28.5% of the relevant year group gained five GCSE grade C, but only 7% did not achieve even a CoE. In fact, the pattern across the three sites is that achievement at one level is inversely related to achievement at another, as has previously been observed at the school level and in comparison between fee-paying and state-funded schools (Gorard 1996$_b$). If the proportion of school-leavers with no qualifications were the benchmark, Blaenau Gwent would have the best educational record in South Wales. The population is uniform with very few Welsh language speakers, 1.2% can speak, read or write Welsh, and almost no residents of recent minority ethnic origin (Welsh Office 1996). 7% of residents had a different address in 1990 to 1991, of whom less than 2% moved from outside the district (see Appendix J).

The local authority runs the largest training centre of its kind in Wales, and there are also five comprehensive schools, three open learning centres, and a campus of Gwent Tertiary College in the authority area.

The three electoral districts

The three electoral divisions chosen to encapsulate the variety of social and educational conditions in the depressed coalfield site of Blaenau Gwent are somewhat varied in size, ranging from Llanhilleth with a population of 5,372 to Cwm with 3,760 (see Appendix K). Beaufort is in the north of the authority, almost outside the valleys and near the heads of the valleys road just south of the Brecon Beacons and Black Mountains and running west through Merthyr Tydfil to Neath. It is semi-rural in nature with scattered housing and small villages with the addition of recent industrial and business parks. While most of the occupations in Beaufort are classified as manual, just as they are in all of Blaenau Gwent, few of these jobs are unskilled or even partly-skilled. In addition, Beaufort has many more professionals than the other two districts and 7% of those economically active are educated to degree level. Even so, only 50% of the adult population is in a full-time job.

Cwm is directly in the middle of the valley, a traditional mining village with one long terraced street on the floor of the valley. Of those economically active, 4% are educated to degree level. Cwm has the lowest unemployment of the three districts, and most of those employed are in manual occupations. Llanhilleth is another mining village, near the Newport end of the valley, with more varied housing than Cym, spread in tiers up the side of the valley. Of those economically active, 1% are educated to degree level, the lowest figure of all nine districts taking part in the survey. The district has a low rate of economic activity, and most of those in jobs have occupations that are manual, many of them unskilled or partly skilled. In all three districts the major industrial classification for womans' jobs is "other services", for men it is still mining-related in Cwm and Beaufort, but metal-related in Llanhilleth.

ACKNOWLEDGEMENTS

The ESRC (Grant Number L123251041), Gwent, Mid Glamorgan and West Wales Training Enterprise Councils for funding. The families and training providers for participating.

REFERENCES

Aberystwyth Policy Group (1990) **Secondary Education in Rural Wales**, Faculty of Education: University of Wales

Ashton D. and Maguire M. (1986) **Young Adults in the Labour Market**, Department of Employment Research Paper number 85

Ashton D., Spilsbury M. and Maguire M. (1990) **Restructuring the Labour Market: Implications for youth**, Basingstoke: Macmillan

Banks M., Bates I, Breakwell G., Bynner J. and Emler N. (1992) **Careers and Identities**, Milton Keynes: Open University Press

Blaenau Gwent (1996) **Official Guide and Street Plan for the County Borough**

Boddy M., Fielder S. and Rees G. (1990) **The institutional determinants of employers training strategies. Local labour market profiles**, ESRC Project Paper 2

Bynner J. (1989) **Transition to Work: Results from a longitudinal study of young people in four British labour markets**, ESRC 16-19 Initiative occasional papers 4, London: City University

Dale A. and Davies R. (1994) **Analyzing Social and Political Change: A casebook of methods**, London: Sage

DfEe (1995) **Training Statistics 1995**, London: HMSO

Dolton P., Makepeace G. and Treble J. (1994) Measuring the effects of training in the Youth Cohort Study, in McNabb R. and Whitfield K. (Eds.) **The Market for Training**, Aldershot: Avebury, p. 195

Eaton G. (1987) **A History of Neath**, Swansea: Christopher Davies

Fevre R. (1989) **Wales is Closed**, Nottingham: Spokesman

Fevre R. (1996) **Some Sociological Alternatives to Human Capital Theory and their Implications for Research on PoCET**, paper presented at ECER conference, Seville 1996

Gershuny J. and Marsh C. (1994) Unemployment in Work Histories, p.66, in Gallie D., Marsh C. and Vogler C. (Eds.) **Social Change and the Experience of Unemployment**, Oxford: Oxford University Press

Gorard S., Rees G., Fevre R., and Furlong J. (1996) **The social determinants of adult participation in education and training: a framework for analysis**, paper presented to British Educational Research Association Annual Conference, Lancaster - also available on Educati*on-Line*

Gorard S. (1996a) Fee-paying schools in Britain - a peculiarly English phenomenon, in **Educational Review**, 48,1, pp. 89-93

Gorard S. (1996b) **What Are They Paying For? Families using fee-paying schools**, paper presented to British Educational Research Association Annual Conference, Lancaster 1996, also available on Educati*on-Line*

Gwent Training and Enterprise Council Ltd. (1996) **Economic and Labour Market Assessment 1995/96**

Hebron S. and Wykes M. (1991) Gender and Patriarchy in mining communities, in Cross M. and Payne G. (Eds.) **Work and the Enterprise Culture**, London: Falmer, p. 160

Howell R. (1988) **A History of Gwent,** Llandysul, Gomer Press

Humphrys G. (1972) **Industrial Britain. South Wales**, Newton Abbot: David and Charles

Istance D and Rees G (1994) Education and training in Wales: problems and paradoxes revisited, in **Contemporary Wales**, 7, p.7-27

Istance D and Rees G (1995) **Lifelong Learning in Wales: A Programme for Prosperity**, a NIACE Cymru policy discussion paper, Leicester: NIACE

Jenkins E. (1974) **Neath and District. A symposium**, Neath: Elias Jenkins

Jones S. (1991) **The History of Port Talbot**, Port Talbot: Goldleaf

Keen R. (1979) **Old Maesteg and the Bridgend Valleys in Photographs**, Barry: Stewart Williams

Lee E., Forthfer R. and Lorimar R. (1989) **Analyzing Complex Survey Data**, Sage: London

Nash G., Davies T. and Thomas B. (1995) **Workmen's Halls and Institutes**, Cardiff: National Museum of Wales

Maguire M., Maguire S. and Felstead A. (1993) **Factors Influencing Individual Commitment to Lifetime Learning. A literature review**, Employment Department

Mid Glamorgan Training and Enterprise Council Ltd. (1996) **Mid Glamorgan Labour Market Assessment 1995-96**

OPCS (1993) **1991 Census. Report for Wales. Part one**, London: HMSO

Port Talbot (1996) **Port Talbot Official Street Atlas**

Rees G. and Rees T. (1983) Migration, industrial restructuring and class relations: an analysis of South Wales, in Williams G. (Ed.) **Crisis of Economy and Ideology. Essays on Welsh Society 1840-1980**, BSA Sociology of Wales Study Group, p. 103

Rees G. and Thomas M. (1991) From Coalminers to Entrepreneurs? A case study in the Sociology of Re-industrialisation, in Cross M. and Payne G. (Eds.) **Work and the Enterprise Culture**, London: Falmer, p. 57

Roberts K. (1987) ESRC Young People in Society, in **Youth and Policy** 22, pp. 15-24

Roberts K. and Parsell G. (1988) **Opportunity Structures and Career Trajectories from age 16-19**, ESRC 16-19 Initiative occasional papers 1, London: City University

Roberts K. and Parsell G. (1989) **The Stratification of Youth Training**, ESRC 16-19 Initiative occasional papers 11, London: City University

Roker D. (1991) **Gaining the Edge: The education, training and employment of young people in private school**, London: City University

Tremlett N., Park A. and Dundon-Smith D. (1995) **Individual Commitment to Learning. Further findings from the Individuals survey**, Employment Department

Wales 2000 (1996) Wales 2000 Issue 7. Published with the **Western Mail**, July 1996

Webster R. (1990) Education in Wales and the Rebirth of a Nation, in **History of Education**, 19, 3, pp. 183-194

Welsh Office (1995a) **Statistics of Education and Training in Wales: Schools, No. 2**, Cardiff: HMSO

Welsh Office (1995b) **Wales in Figures**, Cardiff: Welsh Office Publications Unit

Welsh Office (1995c) **1994/95 Welsh Training and Education Survey**, Cardiff: Welsh Office

Welsh Office (1996) **Digest of Welsh Local Area Statistics 1996**, Cardiff: HMSO

West Wales Training and Enterprise Council Ltd. (1996) **West Wales Annual Labour Market Assessment 1995/96**

Williams G. (1983) **Crisis of Economy and Ideology. Essays on Welsh Society 1840-1980**, BSA Sociology of Wales Study Group

Williams J. (1983) The economic structure of Wales since 1850, in Williams G. (Ed.) **Crisis of Economy and Ideology. Essays on Welsh Society 1840-1980**, BSA Sociology of Wales Study Group, p. 35

APPENDIX A - Background data for Wales

Figure 1 - Industrial Classifications in Wales (Welsh Office 1995c)

	Males	Females
Agriculture, forestry, fishing	5	2
Manufacturing	31	12
Construction	12	2
Distribution, retail, catering	23	25
Finance, real estate	8	9
Public services	16	42
Other	7	8

Figure 2 - Occupational Classifications in Wales (Welsh Office 1996)

	Total	Males
Managers/admin.	15	17
Professional	9	11
Associate	10	7
Clerical	14	5
Craft	11	22
Personal	10	6
Sales	9	5
Plant operatives	14	18
Other	8	8

APPENDIX B - Background data for Industrial South Wales

Figure 1 - Social class of residents

Class	UK 1981	S Wales 1981	UK 1991	S Wales 1991
I	5	3	5	3
II	19	14	21	16
IIIN	9	7	9	7
IIIM	26	29	20	21
IV	12	13	10	10
V	4	5	3	4
Other	25	29	31	39

including armed forces, government scheme, retired, inactive

Figure 2 - Population

UK 1981	S Wales 1981	UK 1991	S Wales 1991
52760331	1327111	54888844	1337741

Figure 3 - Proportion of population with a higher qualification

UK 1981	S Wales 1981	UK 1991	S Wales 1991
7	6	10	8

Figure 4 - Economic position by gender

	UK 1981 Males	Females	S Wales 1981 Males	Females	UK 1991 Males	Females	S Wales 1991 Males	Females
Active	78	46	74	42	73	50	66	44
Unemployed	8	3	10	4	8	3	10	3
Employed FT					50	26	46	23
Employed PT					2	16	2	17
Self-employed with employees					4	1	3	1
Self-employed w/o employees					8	2	6	1
Govt scheme					1	1	1	0

Figure 5 - Occupational classifications

Occupation	UK 1981 Male	Female	S Wales 1981 Male	Female
Agriculture	3	1	1	0
Energy	4	1	12	1
Manufacturing	32	20	35	21
Construction	11	1	11	4
Distribution	15	25	12	26
Transport	9	3	8	3
Other	25	48	20	47

Occupation	UK 1991 Male	Female	S Wales 1991 Male	Female
Agriculture	3	1	1	0
Energy	3	1	4	1
Mining	4	2	8	2
Metal	13	4	13	8
Other manuf	9	8	10	9
Construction	12	2	14	1
Distribution	17	24	15	24
Transport	9	3	7	2
Banking	11	13	7	8
Other	18	41	2	44

APPENDIX C - Background data for Mid Glamorgan area

Figure 1 - Highest qualification equivalent as a percentage of the workforce (Welsh Office 1995c)

	Male	Female	Average	Wales
Degree	6	5	6	9
HE < degree	8	8	8	9
A level	30	17	24	24
GCSE	15	23	19	19
Other	14	13	14	12
None	27	34	30	27

Figure 2 - Employment Mode and Industrial Classification, Mid Glamorgan (Welsh Office 1995c)

Status	Mid Glamorgan	Wales
Full-time	68	60
Part-time	21	22
Full-time self-employed	10	15
Part-time self-employed	2	2
Agriculture	1	3
Manufacturing	29	23
Construction	7	7
Distribution	21	24
Finance	10	9
Public services	26	27
Other	6	7

APPENDIX D - Background data for Unitary Authority of Bridgend

Figure 1 - Population of Bridgend/Ogwr

Resident 1981	Resident 1991
126809	132442

Figure 2 - Percentage of population with a higher qualification

Bridgend 1981	Bridgend 1991
7	9

Figure 3 - Social Class of Bridgend/Ogwr employed residents

Class	1981	1991
I	5	6
II	22	27
IIIN	10	10
IIIM	42	34
IV	15	16
V	5	5

Figure 4 - Economic position of residents of Bridgend/Ogwr by gender

Position	Males 1981	Males 1991	Females 1981	Females 1991
Active	74	67	42	45
Employed FT		47		24
Employed PT		1		15
Self-employed		8		3
Govt scheme		1		0
Unemployed	9	9	3	3

Figure 5 - Occupational classifications for Bridgend/Ogwr 1981

Occupation	Male 1981	Female 1981
Agriculture	1	0
Energy	13	1
Manufacturing	32	21
Construction	13	2
Distribution	12	24
Transport	7	2
Other	20	49

Figure 6 - Occupational classifications for Bridgend/Ogwr 1991

Occupation	Male 1991	Female 1991
Agriculture	1	0
Energy	3	0
Mining	7	2
Metal	17	9
Other manuf	12	10
Construction	14	1
Distribution	14	23
Transport	6	2
Banking	7	7
Other services	18	44

APPENDIX E - Background data for Coity, Newcastle and St Bride's

Figure 1 - Population of the three Mid Glamorgan districts (1991)

	Coity Higher	Newcastle	St Bride's Minor
Houses	1850	2261	2106
Persons	5224	5404	5755

Figure 2 - Percentage of population qualified beyond A level or equivalent (1991)

Coity Higher	Newcastle	St Bride's Minor
16	12	5

Figure 3 - Social class of the three Mid Glamorgan districts(1991)

Class	Coity Higher	Newcastle	St Bride's Minor
I	10	9	2
II	45	33	20
IIIN	10	12	12
IIIM	21	31	42
IV	16	12	12
V	1	3	12

Figure 4 - Migrancy in the three Mid Glamorgan districts 1990-1991

	Coity Higher	Newcastle	St Bride's Minor
Total moves	7	8	8
From outside	3	2	2

Figure 5 - Economic position by gender in the three Mid Glamorgan districts (1991)

	Coity		**Newcastle**		**St Bride's**	
Position	Male	Female	Male	Female	Male	Female
Active	67	50	63	44	71	51
FT employ	51	29	44	23	51	27
PT employ	1	16	2	15	1	17
Self-employ	8	3	8	3	7	2
Scheme	0	0	1	0	1	1
Unemploy	5	3	8	3	11	3

Figure 6 - Occupation in the three Mid Glamorgan districts (1991)

	Colty		Newcastle		St Bride's	
Occupation	Male	Female	Male	Female	Male	Female
Agriculture	0	0	0	0	1	1
Energy	5	0	6	0	2	0
Mining	7	0	6	1	6	1
Metal	14	3	8	1	14	11
Other manu	12	1	10	5	13	11
Construct'n	7	2	15	1	13	0
Distribution	10	22	16	24	16	22
Transport	3	1	8	1	8	1
Banking	13	14	5	15	11	4
Other serv	29	57	26	52	18	47

APPENDIX F - Background data for West Glamorgan area

Figure 1 - Highest qualification held by persons of working age (Welsh Office 1995c)

	Male	Female	Total	Wales
Degree	13	8	10	9
HE < degree	9	10	9	9
A level	29	18	24	24
GCSE	15	22	18	19
Other	11	13	12	12
None	23	30	26	27

Figure 2 - Employment Mode and Industrial Classification in West Wales (Welsh Office 1995c)

Status	West Wales	Wales
Full-time	60	60
Part-time	22	22
Full-time self-employed	15	15
Part-time self-employed	2	2
Agriculture	6	3
Manufacturing	20	23
Construction	8	7
Distribution	24	24
Finance	7	9
Public services	28	27
Other	7	7

APPENDIX G - Background data for Unitary Authority of Neath Port Talbot

Figure 1 - Population of Neath Port Talbot

Resident 1981	Resident 1991
119746	116423

Figure 2 - Percentage of population with a higher qualification

Neath Port Talbot 1981	Neath Port Talbot 1991
5	6

Figure 3 - Social Class of Neath Port Talbot employed residents

Class	1981	1991
I	4	4
II	13	21
IIIN	9	10
IIIM	42	36
IV	22	21
V	10	7

Figure 4 - Economic position of residents of Neath Port Talbot by gender

Position	Males 1981	Males 1991	Females 1981	Females 1991
Active	73	63	40	41
Employed FT		45		21
Employed PT		1		14
Self-employed		7		2
Govt scheme		2		1
Unemployed	11	9	4	3

Figure 5 - Occupational classifications for Neath Port Talbot 1981

Occupation	Male 1981	Female 1981
Agriculture	1	0
Energy	12	1
Manufacturing	42	21
Construction	12	1
Distribution	10	27
Transport	7	2
Other	15	46

Figure 6 - Occupational classifications for Neath Port Talbot 1991

Occupation	Male 1991	Female 1991
Agriculture	1	0
Energy	7	1
Mining	11	1
Metal	21	9
Other manuf	6	9
Construction	13	1
Distribution	13	24
Transport	8	2
Banking	5	7
Other services	14	44

APPENDIX H - Background data for Cimla, Bryn Cymavon, Sandfields West

Figure 1 - Population of the three West Glamorgan districts (1991)

	Cimla	Bryn Cymavon	Sandfields West
Houses	1545	2510	2662
Persons	4455	6243	6975

Figure 2 - Percentage of population qualified beyond A level or equivalent (1991)

Cimla	Bryn Cymavon	Sandfields West
10	6	2

Figure 3 - Social class of the three West Glamorgan districts (1991)

Class	Cimla	Bryn Cymavon	Sandfields West
I	5	6	1
II	29	22	11
IIIN	16	10	7
IIIM	32	35	41
IV	15	17	21
V	3	9	18

Figure 4 - Migrancy in the three West Glamorgan districts 1990-1991

	Cimla	Bryn Cymavon	Sandfields West
Total moves	6	8	5
From outside	2	2	1

Figure 5 - Economic position by gender in the three West Glamorgan districts (1991)

	Cimla		**Bryn**		**Sandfields**	
Position	Male	Female	Male	Female	Male	Female
Active	77	69	63	39	56	36
FT employ	62	32	46	22	38	19
PT employ	2	22	1	13	1	13
Self-employ	7	2	6	2	4	0
Scheme	1	1	1	1	2	1
Unemploy	6	3	8	2	12	3

Figure 6 - Occupation in the three West Glamorgan districts (1991)

	Cimla		Bryn		Sandfields	
Occupation	Male	Female	Male	Female	Male	Female
Agriculture	0	0	3	0	0	0
Energy	10	0	3	0	1	2
Mining	10	2	17	3	20	0
Metal	15	4	16	8	27	11
Other manu	6	4	8	6	5	18
Construct'n	10	1	10	1	19	0
Distribution	14	29	10	33	6	27
Transport	10	2	8	1	9	1
Banking	7	9	7	10	3	6
Other serv	20	48	15	37	9	33

APPENDIX I - Background data for Gwent Research Site

Figure 1 - Highest qualification held by persons of working age (Welsh Office 1995c)

	Gwent	Wales
Degree	8.4	9.2
HE < degree	7.7	8.6
A level	21.8	24.3
GCSE	19.0	18.9
Other	13.2	12.0
None	29.9	27.1

Figure 2 - Employment Mode and Industrial Classification in Gwent (Welsh Office 1995c)

Status	Gwent	Wales
Full-time	67	60 check ???
Part-time	22	22
Full-time self-employed	9	15
Part-time self-employed	2	2
Agriculture	2	3
Manufacturing	32	23
Construction	6	7
Distribution	23	24
Finance	6	9
Public services	25	27
Other	6	7

APPENDIX J - Background data for Unitary Authority of Blaenau Gwent

Figure 1 - Population of Blaenau Gwent

Resident 1981	Resident 1991
78864	72254

Figure 2 - Percentage of population with a higher qualification

1981	1991
4	5

Figure 3 - Social Class of Blaenau Gwent employed residents

Class	1981	1991
I	2	3
II	14	20
IIIN	8	9
IIIM	44	38
IV	22	21
V	9	9

Figure 4 - Economic position of residents of Blaenau Gwent by gender

Position	Males 1981	Males 1991	Females 1981	Females 1991
Active	71	64	39	41
Employed FT		45		22
Employed PT		1		14
Self-employed		6		2
Govt scheme		2		1
Unemployed	13	10	4	3

Figure 5 - Occupational classifications for Blaenau Gwent (1981)

Occupation	Male 1981	Female 1981
Agriculture	1	0
Energy	19	1
Manufacturing	42	29
Construction	11	1
Distribution	8	26
Transport	5	2
Other	15	40

Figure 6 - Occupational classifications for Blaenau Gwent (1991)

Occupation	Male	Female
Agriculture	1	0
Energy	3	0
Mining	17	2
Metal	21	11
Other manuf	11	12
Construction	12	1
Distribution	11	22
Transport	6	3
Banking	4	6
Other services	14	42

APPENDIX K - Background data for Beaufort, Cwm and Llanhilleth

Figure 1 - Population of the three Gwent districts (1991)

	Beaufort	Cym	Llanhilleth
Houses	1524	1503	2151
Persons	4026	3760	5372

Figure 2 - Percentage of active population qualified beyond A level or equivalent (1991)

Beaufort	Cym	Llanhilleth
7	4	1

Figure 3 - Social class of the three Gwent districts (1991)

Class	Beaufort	Cym	Llanhilleth
I	8	1	0
II	17	18	12
IIIN	8	8	9
IIIM	44	37	36
IV	14	22	28
V	9	14	16

Figure 4 - Migrancy in the three Gwent districts (1990-1991)

	Beaufort	Cym	Llanhilleth
Total moves	8	8	7
From outside	1	2	2

Figure 5 - Economic position by gender in the three Gwent districts (1991)

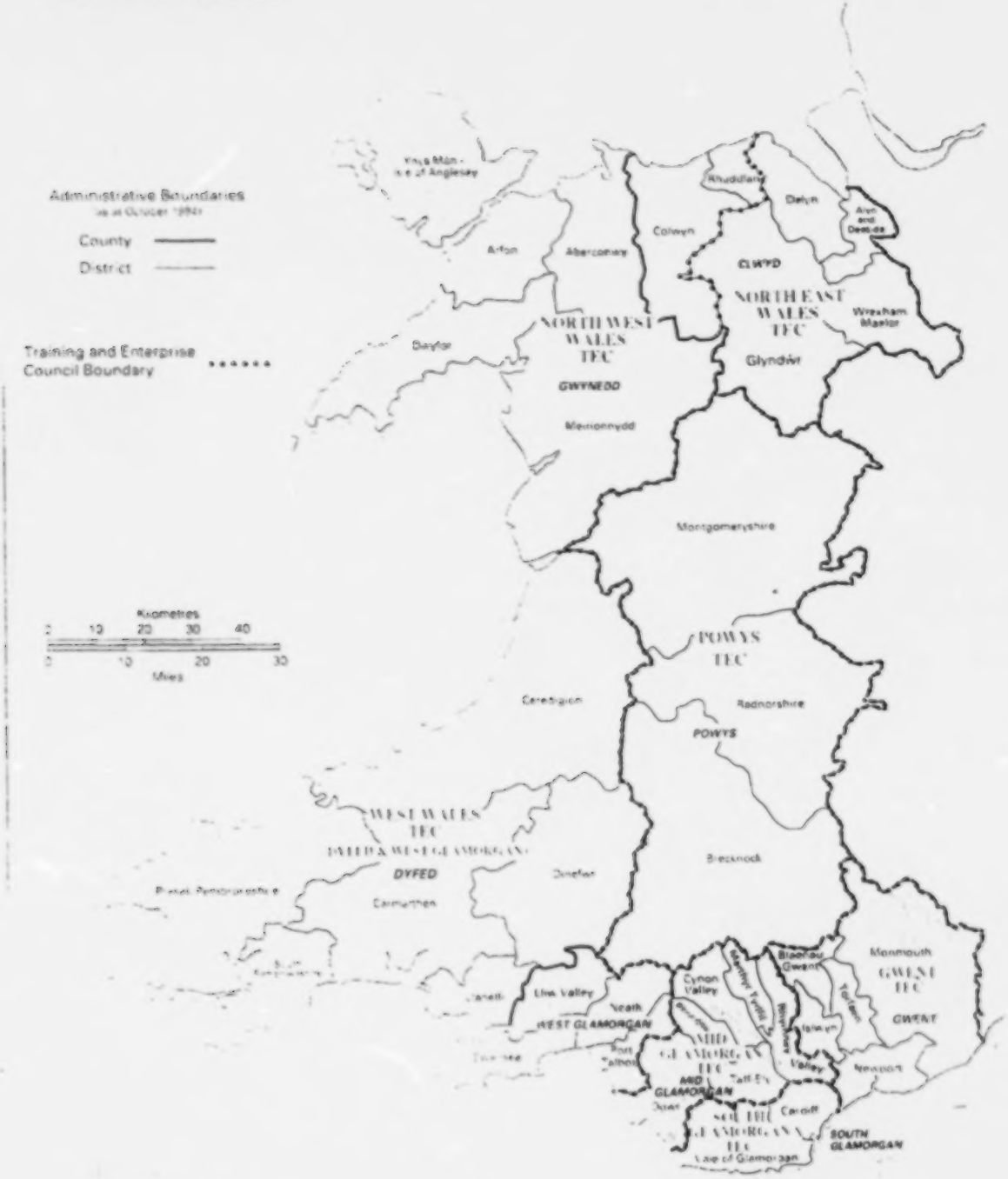
	Beaufort		**Cym**		**Llanhilleth**	
Position	Male	Female	Male	Female	Male	Female
Active	67	48	63	39	61	36
FT employ	50	27	46	20	43	20
PT employ	1	16	1	14	1	11
Self-employ	4	2	6	2	5	2
Scheme	1	1	2	1	2	1
Unemploy	10	3	9	3	11	3

Figure 6 - Occupations in the three Gwent districts (1991)

	Beaufort		Cym		Llanhilleth	
Occupation	Male	Female	Male	Female	Male	Female
Agriculture	0	0	1	0	1	0
Energy	4	0	7	0	0	0
Mining	27	2	26	0	11	4
Metal	16	11	19	9	26	20
Other manu	9	11	16	16	14	9
Construct'n	18	0	6	0	14	0
Distribution	5	22	12	19	17	25
Transport	1	0	3	5	9	6
Banking	5	14	1	2	2	4
Other serv	14	40	10	47	6	33

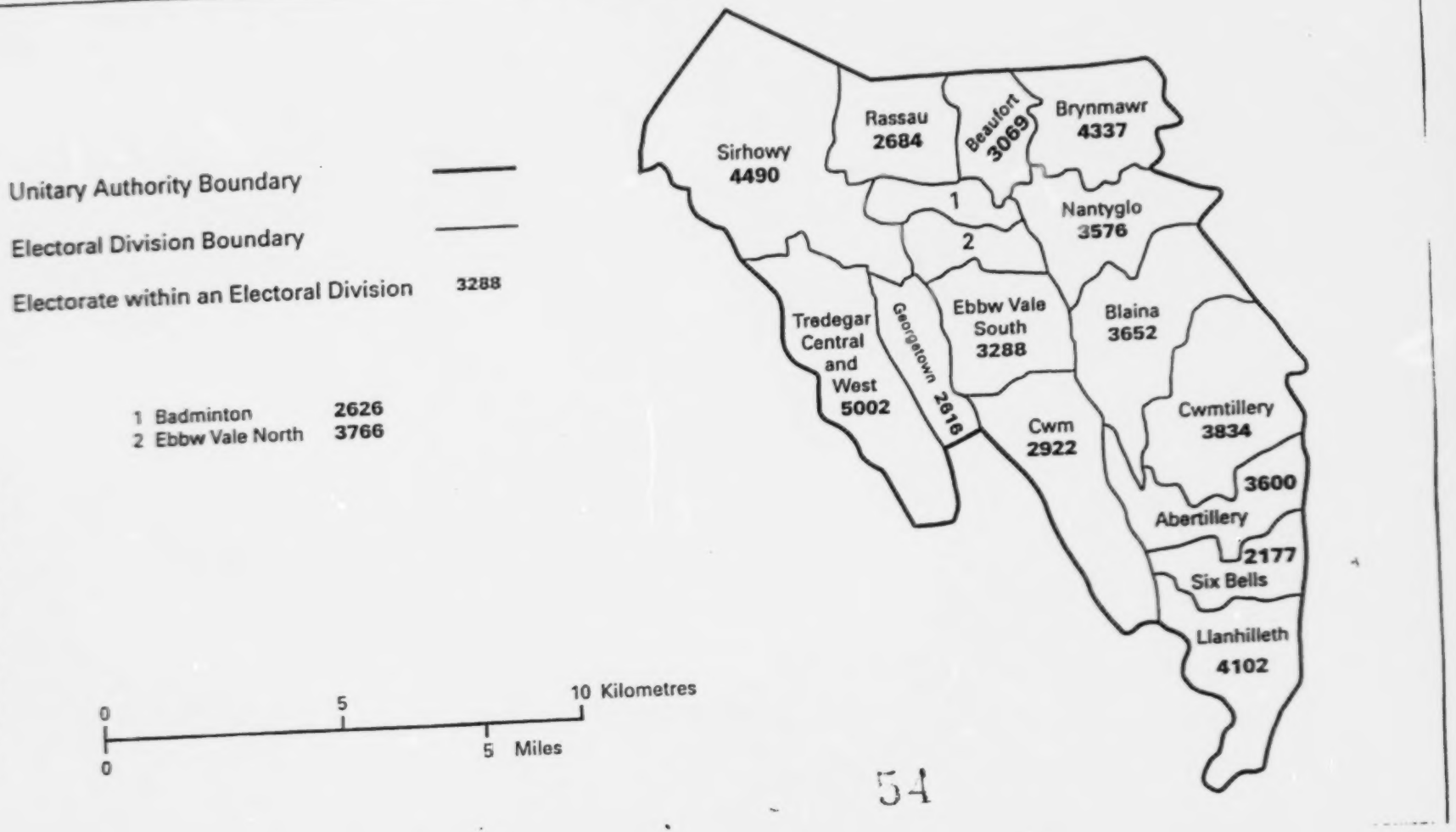
APPENDIX L - Maps of the research areas

WALES Training and Enterprise Councils



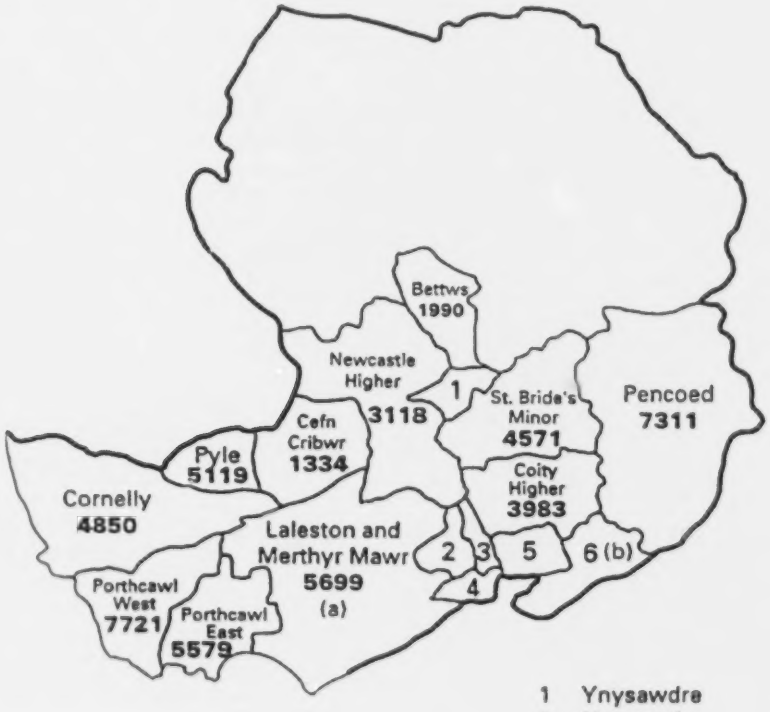
BEST COPY AVAILABLE

Blaenau Gwent - New Electoral Divisions



54
BEST COPY AVAILABLE

Bridgend - New Electoral Divisions



1	Ynysawdre	**2096**
2	Newcastle	**4389**
3	Morfa	**3774**
4	Oldcastle	**3323**
5	Brackla	**6380**
6	Coychurch Lower	**1113**

Unitary Authority Boundary ——

Electoral Division Boundary ——

Electorate within an Electoral Division **5579**

(a) Laleston Ward and Community of Merthyr Mawr
(b) Coychurch Lower Community area only

0 5 10 Kilometres
0 5 Miles

BEST COPY AVAILABLE

Neath Port Talbot - New Electoral Divisions

1	Gwaun-Cae-Gurwen	2265
2	Lower Brynamman	1053
3	Godre'r graig	1218
4	Trebanos	1145
5	Cadoxton	1123
6	Bryn-côch South	2845
7	Coedffranc North	1898
8	Coedffranc Central	3166
9	Neath North	3381
10	Neath South	3622
11	Briton Ferry West	2390
12	Sandfields West	5314
13	Aberavon	4449
14	Sandfields East	4994
15	Port Talbot	4391

Unitary Authority Boundary

Electoral Division Boundary

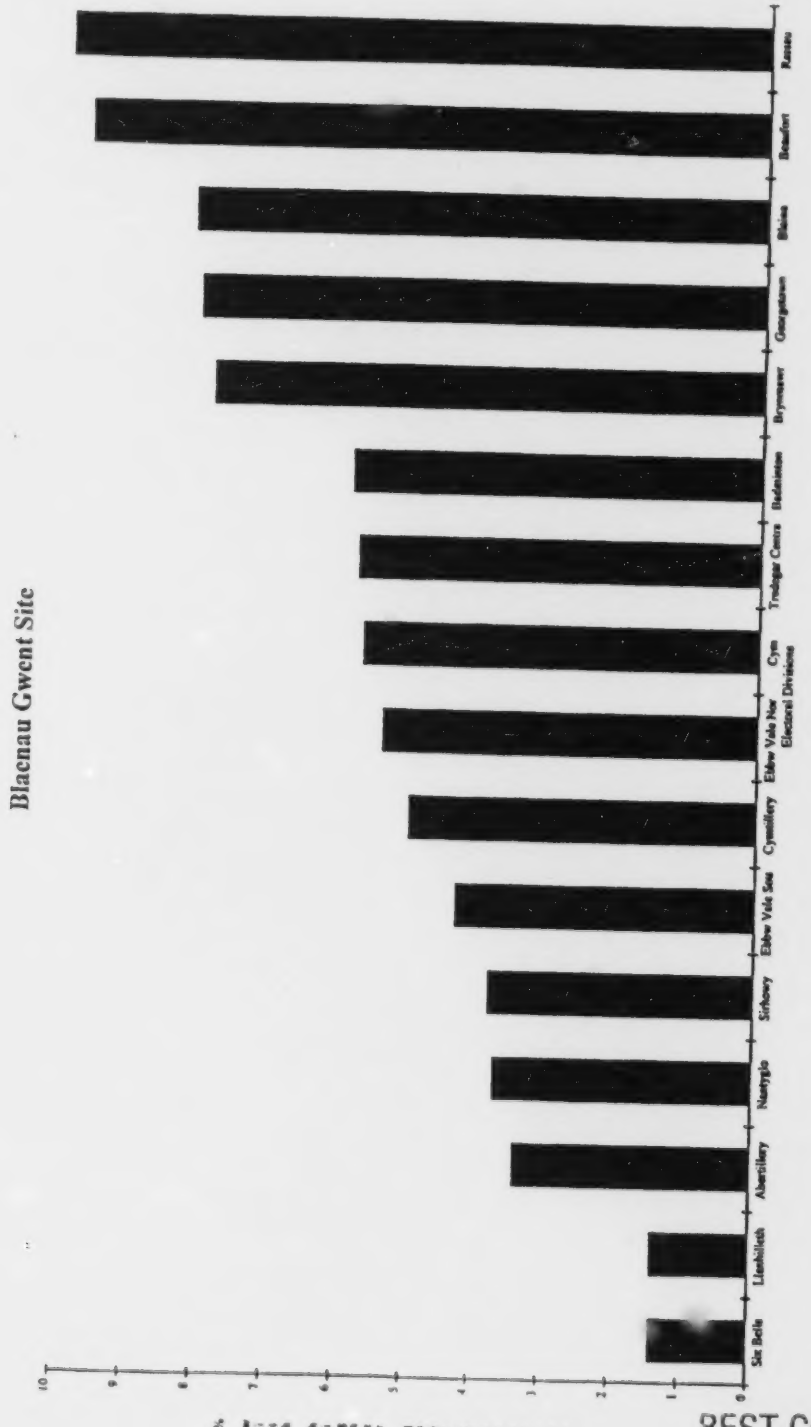
Existing District Boundary

Electorate within an Electoral Division 1659

0 5 10 15 Kilometres

0 5 10 Miles

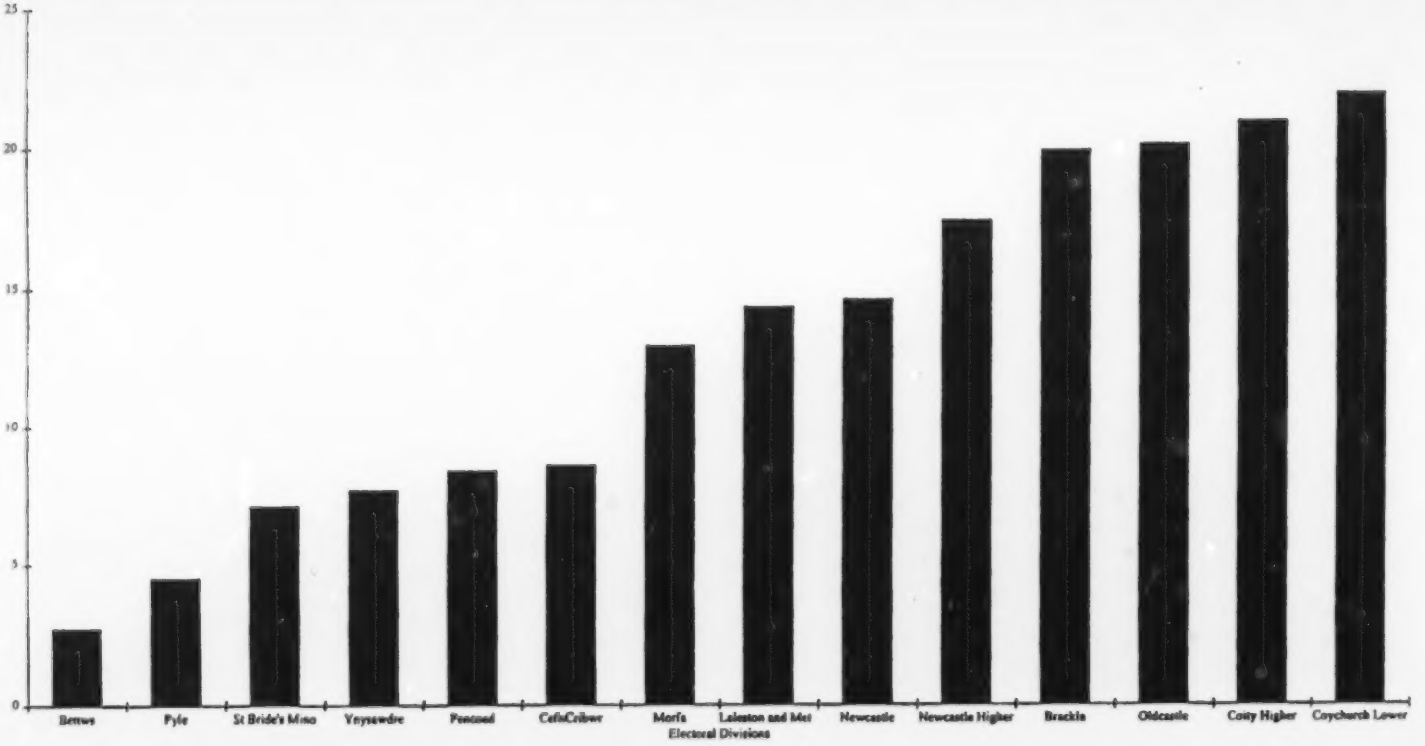
APPENDIX M - Educational profiles of the three research sites



BEST COPY AVAILABLE

57

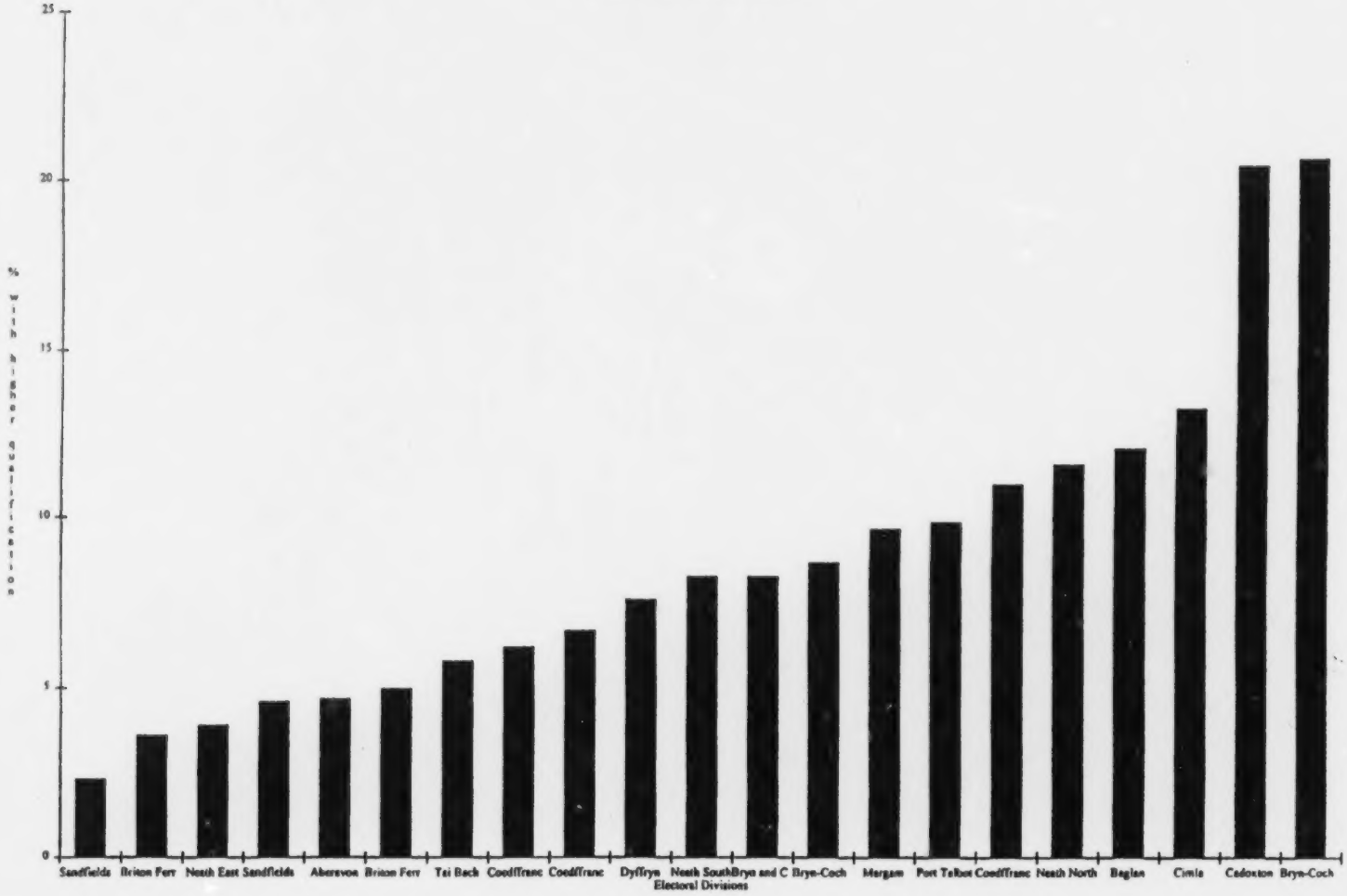
Bridgend Site



58

BEST COPY AVAILABLE

Neath Port Talbot Site



BEST COPY AVAILABLE

ISBN 1 872330 02 9

CE

PATTERNS OF PARTICIPATION IN ADULT EDUCATION AND TRAINING

A Cardiff and Bristol University ESRC- funded Learning Society Project

WORKING PAPER 2

OUTLINE METHODOLOGY OF THE STUDY

Stephen Gorard, Gareth Rees, John Furlong, Ralph Fevre

1997

SCHOOL OF EDUCATION

Cardiff University of Wales

61

LIST OF CONTENTS

	PAGE NUMBER
Introduction	1
1. The sample	1
Weighting the sample	4
2. The survey	5
The instrument	7
Data collection	9
The pilot study	9
Coding	11
3. Analysis	12
Factor analysis	14
Causation	17
Temporal sequences	20
Standard regression	22
Logistic regression	24
Loglinear models	26
Conclusion	28
Acknowledgements	29
References	29
Appendix	33

ISBN 1 872330 03 7

PATTERNS OF PARTICIPATION IN ADULT EDUCATION AND TRAINING

A Cardiff and Bristol University ESRC-funded Learning Society Project

WORKING PAPER 2

OUTLINE METHODOLOGY OF THE STUDY

Stephen Gorard, Gareth Rees, John Furlong and Ralph Fevre

INTRODUCTION

The project, funded by the ESRC as part of the Learning Society Programme, is a regional study in industrial South Wales of the determinants of participation and non-participation in post-compulsory education and training, with special reference to processes of change in the patterns of these determinants over time and to variations between geographical areas. The study combines contextual analysis of secondary data about education and training providers with a regional study of several generations of families in South Wales via survey, semi-structured interviews and taped oral histories. The background to the study is further described in Gorard et al. (1996). The study took place in three markedly different centres in industrial South Wales during 1996 and 1997. This paper describes the methods used in the survey stage of the research.

1. THE SAMPLE

As in previous studies, the decision as to which geographical areas to use for the sample was made partly on strategic grounds (Banks et al. 1992) and partly on the basis of the population characteristics in each area. Although it is normal for a study of this size to select a diverse and well-spread sample which can then be divided into gender or class categories, since the differences in local effects are so varied in the UK a more focused study was required to make these local effects explicit. Three sites were chosen to represent the range of experiences in industrial South Wales. Bridgend could be briefly characterised as an expanding town, Neath Port Talbot as an established manufacturing conurbation, and Blaenau Gwent as a depressed coalfield valley. Three electoral divisions were selected within each site to reflect the range and diversity of their social and educational conditions, as described in Gorard (1997a). Secondary analysis was used to characterise the nine wards, the three sites, and South Wales in terms of a

range of social and educational measures capable of disaggregation at this level. This data provided the sampling frame for the survey and part of the context for the primary analysis.

Households in each division were identified from the electoral registers, which were the most accurate and up-to-date listings available of individuals and their addresses. These lists were appropriate since the target population for the first wave was only those adults aged 35 to 64, with children not involved at this stage. In this way the sampling method was similar to that used in some previous research (cf. Gershuny and Marsh 1994). Lists from organisations such as colleges and companies could not be used since the study was unusual in trying to contact non-participants in education and employment as well as participants (Whitfield and Bourlakis 1991).

Systematic sampling, selecting every *n*th element after a random start, was an appropriate method to use but to avoid problems of periodicity and special methods of variance estimation, repeated systematic sampling was used. The sample was also stratified in an attempt to reduce the sampling variance and to ensure sufficient cases in certain categories, but the stratification was proportionate i.e. the sampling fraction for each strata was uniform within wards. Stratification was appropriate as the sample to population ratio was large, and the population was well characterised (Lehtonen and Pahkinen 1995). Four points were chosen at random in each electoral division and every *n*th household after those points was selected for inclusion in the survey as a primary respondent with the ten subsequent addresses used as the reserve respondents. The sampling fraction $1/n$ was similar for each division but was determined by the precise number of electors (Gorard 1997a). A quota was devised such that half of the respondents in each electoral division were male and half female. For the first wave, one third were aged 35 to 44, one third were 45 to 54 and the remainder 55 to 64.

The primary address was visited three times at different times of the day in successive weeks until contact was made with one of the householders. In a house with two or more householders, either was interviewed, according to quota and as convenient. If the house was clearly empty (having been "gutted" by builders for example) or all of the householders were out-of-quota, the first house on the

reserve list became a new primary and the process started again. If the primary householder refused to take part or was not contacted after three calls, the first house on the reserve list was used instead, followed, if necessary, by the second reserve and so on. In the latter cases the response was recorded as a reserve for accounting purposes. This procedure of repeated systematic sampling produced a set of respondents stratified by age and gender within electoral divisions chosen to represent the educational and socio-economic diversity of the three research sites, which were themselves selected as representatives of the range of socio-economic experiences in industrial South Wales over the past 50 years (Gorard 1997a).

The second wave of the survey was based in the same areas as the first and was also stratified by gender and age. Half of the respondents in the second wave were aged 15-24 and half were 25-34. These became the two younger cohorts, so that the sample as a whole represented education and training experiences throughout working life from just before finishing compulsory education to retirement. However, in order that the study could also examine family influences more fully the two younger cohorts were selected by repeated systematic sampling from the children of those in the first wave. Several of the respondents in the second wave therefore no longer lived in South Wales or even the UK. This second stage was similar in many ways to the method of boosting a sample in order to obtain a higher proportion of respondents with certain characteristics (SCELI 1991, Gershuny and Marsh 1994, Park and Tremlett 1995). The survey was designed to include at least 800 respondents in the first wave and at least 200 in the second. From this grand total of over 1,000, between 150 and 200 were selected strategically for in-depth interviews. There were, in addition, 20 in-depth interviews with local training providers who had some knowledge of changes in the structure of opportunities for formal learning over the past 50 years. In this way, the study involved a large-scale survey comparable in size to those of Deloitte Haskins and Sells (1989), Davies (1994), or Park (1994), or to one research site in larger studies (e.g. SCELI 1991, Banks et al. 1992), as well as a more qualitative segment comparable to previous studies of training providers, such as Pettigrew et al. (1989), and considerably larger than many studies of participants in training (e.g. FEU 1993, Hand et al. 1994, Taylor and Spencer 1994).

Weighting the sample

In using sites and wards as administrative units to select the sample, the survey can be described as "complex", so that some standard statistical procedures might not be appropriate without weighting (Lee et al. 1989). Social science sample selection is seldom simply random and sample designs have generally run ahead of analytic solutions, partly because so many surveys use only descriptive statistics. However, any deviation from the random "ideal" can lead to biased results for almost all techniques and this applies to any survey with a restriction other than simple random sampling with replacement (Lee et al. 1989). There was little value for this study in using simple random sampling of households within wards to avoid weighting, since the survey was already complex, and some form of post-stratification weights may have been needed anyway.

Weighting was possible since the entire sampling frame for the survey at each stage of this multi-stage design was known, at least in principle, and each element had a known non-zero chance of being included in the sample (Lee et al. 1989) since a record was kept of the probabilities of inclusion at each stage of the design. However, the sample was automatically self-weighting at the ward level since the probability of selection within a ward was proportionate to the size of the ward. Similarly within each ward, although each household of whatever size had an equal chance of selection, the sample was of householders only and stratified by gender, therefore it was not necessary to correct for the probability of selection in terms of the number of people in the house (Lee et al. 1989), as has been done elsewhere, for example by using the Kisch procedure and adjustments (SCELI 1991, Gershuny and Marsh 1994).

Weights are used in this survey for two main reasons: to attempt a correction for missing values and to form a more reliable estimate of the population characteristics for industrial South Wales, since even a self-weighting sample can have non-response and design errors. Where a questionnaire was incomplete or had other missing values, a simple method of imputing the missing value was devised if possible (Lehtonen and Pahkinen 1995, Lee et al. 1989), such as using the mean of all other cases. To attempt a correction for bias in the sample through self-selection, where this was related to demographic composition, and at the same time expand the data to fit the background characteristics of industrial South Wales better (Gorard 1997a), the responses were reweighted via a post-

stratification adjustment factor (Lehtonen and Pahkinen 1995, Lee et al. 1989). This factor adjusted for the different response rates within categories and adjusted demographic variables to those of the known population. The adjustment factor is the population distribution divided by the sample distribution for each demographic subgroup. For example, if the proportion of men aged 35-44 in South Wales is 12%, and the proportion in the sample was 10%, the factor is 1.2 (Lee et al. 1989). Weighting and imputing were only used for the production of descriptive statistics and the estimation of current population parameters. The proposed analysis of the determinants of participation in learning covers 50 years from 1945 to 1996, therefore no one set of population characteristics from a snapshot date would be appropriate to use as the basis of weights for that whole period.

2. THE SURVEY

Since this project examines changes in learning histories and opportunities over a period of at least 50 years, a much longer time period than most studies (Whitfield and Bourlakis 1991), a longitudinal study allowing prolonged study of the lives of respondents would have had many attractions. Data from such a study could be richer, may be more accurate and would certainly help to understand the process of change. However, it would also be expensive and time-consuming, so that the decision to reject a longitudinal study was effectively made by the funders. A longitudinal study would also have entailed many compromises. It would lead to complex statistical problems, such that longitudinal data is often collapsed into a format of one or more cross-sections for analysis anyway (Crouchley 1987). Long term studies also suffer from respondent attrition, so that even the best ones may end up with an overall response rate which clearly suggests bias through self-selection in the sample (Dolton et al. 1994). For example, Banks et al. (1992) had response rates of between 60% and 70% for the first sweep of their study, but if this response rate was similar on each occasion that they attempted to contact the original respondents in successive years, the overall response rate for the third sweep could easily be less than 25%. The respondents in the Banks study became proportionately more middle-class in each wave. Similarly only 45% of the respondents in the Youth Cohort Study took part in all three sweeps (Whitfield and Bourlakis 1991). In addition to respondent mortality, longer studies also face the threat to internal validity coming from the necessity to test and re-test the same individuals (Hagenaars 1990). One way

around this is to use a trend design (Hagenaars 1990) collecting data from different groups for each sweep, but such a design is not possible for this study which concerns the changes in individual's lives. Also, the second sample in a trend design is not from the same population as the first in the statistical sense and this causes problems in looking for changes in population parameters over time.

Both retrospective and longitudinal studies face problems of comparability over time (Glenn 1977). For example, the modes and titles of certified public examinations change over time and even where equivalences between them are established it is not clear that their value-in-exchange remains constant. An 'A' level may have meant a lot more in 1970 than in 1990, not because it was any harder to attain, but simply because there were fewer of them. However, such considerations are more problematic for a long term study since the instrument to be used for all sweeps has to be designed before the changes that it needs to encompass, or else has to be changed between sweeps, exacerbating the comparability problem. A retrospective study has the advantage of hindsight. A retrospective, as opposed to a simple cross-sectional study also avoids many of the other problems noted above. Retrospective learning histories, such as the National Training Survey 1975/76 (Greenhalgh and Stewart 1987), and employment histories, such as the 1984 Women and Employment Survey (Martin and Roberts 1984), are much used by economists (McNabb and Whitfield 1994). They are not, of course, immune to criticism since a wide range of life variables may be difficult for the respondents to recall. Among these variables are attitudes, which are notoriously unreliable post hoc, and some figures such as income and health measurements (Dale and Davies 1994). However, such variables are in the main avoided by the present study (see Appendix).

Even so, if educational and occupational data is seen as being continuous in nature, a snap-shot study may well bias the reporting due to the selective and fallible nature of memory (Hagenaars 1990). However, longitudinal design does not in itself produce more reliable reporting than simple recall. For example, Dolton et al. (1994) found a substantial degree of month-to-month inertia in the economic activity states of their respondents in each sweep, but found considerable change at the join between successive questionnaires. There were often different states reported at the end of one sweep to those reported at the start of another, even though they were meant to be contiguous. Similarly there

were spikes of changes every six months in the reports, with more uniform rates of change in between, suggesting that six months is a psychologically convenient unit of memory for respondents. In general any methodological problems raised by the use of retrospective histories can be overcome by a series of measures including triangulation of methods, consideration of omitted variables, care in explaining results and parsimony at all times (Hagenaars 1990).

The instrument

The outline design for the instrument used in the survey was based upon three assumptions: that it would in involve a retrospective history of education and training for adults up to the age of 65, that it would be administered face-to-face by professional interviewers employed by the project for the purpose, and that, unlike some surveys (Banks et al. 1992), the same instrument would be used in all three research sites and for both generations. The potential length of the work and training histories for those individuals nearing retirement meant that the questionnaire could not gather much else besides, without becoming unwieldy and endangering the response rate or causing fatigue. The instrument was budgeted to last for an average of 45 minutes. After pre-piloting it was decided to use any spare time to ask for abbreviated histories of other family members, rather than adding attitudinal sections, for example, which could be more easily managed in the later interviews and are anyway the weakest part of retrospective work (see above). Although the study is concerned with post-compulsory education, to start the history at age 16 for example would screen out earlier experiences - "left-censoring" (Tuma 1994) - and deny their importance in the creation of training trajectories (forthcoming paper). The questionnaire therefore starts at birth and continues through initial education into working life and beyond. Each event in the work history is defined by a change of status, leading to a variable number of fixed length event records per individual (cf. Gershuny and Marsh 1994)

One significant problem was the minimum period of time recorded as an event in the curricula vitae. There is the problem of temporary summer jobs for those intending to return to education and the short mean length of periods of activity for the younger age groups (Dolton et al. 1994). A few weeks employment could be very revealing in the account of a 17 year-old, but to create a 50 year work history for a 65 year-old at that level of detail would be unmanageable. On the other hand, ignoring events of short duration can bias the view of policy-makers,

regarding the nature of unemployment for example (Tuma 1994). Some studies have used the "dipstick" method to find out what respondents were doing at set times without collecting details of interim periods (Bynner 1992). This is done to overcome the problem of timescale, but the method was not used in this study because it was intended to carry out an event history analysis (Tuma 1994). Each account is therefore a collection of pairs of variables recording 'when' and 'what', with the time period being defined by the elapsed time before a change of state. Although a change of employer or job is treated as a change of state, on practical grounds, collections of short-term employment lasting less than six months are defined as a state in themselves, also partly justified by the psychological six month unit of recall described above.

Several previous studies have considered the determinants of continuing initial education into the post-compulsory stage, others have considered the determinants of employment and unemployment, while a further group have examined some general determinants of later training. Questions in the instrument were based on the variables found to be relevant in all of these studies. Such variables are gender (Glenn 1977, Greenhalgh and Mavrotas 1994), race or ethnic background (Mare and Winship 1985, Greenhalgh and Stewart 1987, Whitfield and Bourlakis 1991), disability (Whitfield and Bourlakis 1991), age and age-related indicators such as marital status, number and ages of children, living away from parents home and years since school (Mare and Winship 1985, Greenhalgh and Stewart 1987, Whitfield and Bourlakis 1991, Gershuny and Marsh 1994), parents' and own occupational class history (Gambetta 1987, Gershuny and Marsh 1994), parents' and own economic status (Greenhalgh and Stewart 1987, Main and Shelly 1990), length of employment and industry type (Gershuny and Marsh 1994), and past educational experiences such as attendance at school (Mare and Winship 1985), years in continuous education (Gershuny and Marsh 1994), qualifications (Greenhalgh and Stewart 1987, Main and Shelly 1990), the experience of failure at school (Gambetta 1990), and past training (Main and Shelly 1990). Interacting with and adding to these factors are the attitudes of the prospective trainees, their awareness of opportunities, the local opportunity structure and the policies of the trainers and employers (Dolton et al. 1994). The latter variables are assessed more by other parts of the study than by the survey. They might be seen as lying beneath some of the error components, unexplained variance and residuals encountered in attempting to model the determinants of training statistically.

Data collection

The survey forms were completed by interviewers working for the project. These interviewers were local residents who knew the areas they were working in. They were trained by the researchers and debriefed after completion. The majority of them were also involved in the piloting. The instrument itself was colour-coded so that each skip or branch instruction involved a move to a new coloured section. Sections recording work and educational experiences were repeated, so that interviewers could use as many or as few as needed. Most responses were prompted by a list of possible answers on a set of cards shown to the respondent, although all questions also allowed an open response.

Observation of the interviews by the researchers revealed some of the strength and weaknesses of the methods used. On one occasion, the interviewer had to call at six houses after the primary address in the sample in order to get the first interview of the day. One person was out-of-quota, one refused, two houses had notices asking visitors not to ring bell as the occupant was a shift-worker, and two had been demolished. On another occasion, a friend of the occupant visited with two children just as the partner of the occupant came home from shopping. Their discussions, play and arguments in the other end of the through lounge were an understandable distraction throughout the remainder of the interview.

During the second wave, those children who had moved a long way from the original research site, where their parents had been interviewed, were contacted directly by the researchers. In these cases the interview was most commonly conducted by telephone or post. A few of those living abroad were contacted when they returned to the UK at Christmas 1996.

The pilot study

A pilot study was carried out with a draft version of the questionnaire in June and July 1996, involving 81 cases from all nine wards in all three research sites. The respondents were selected systematically from the electoral register, 27 in each site, and 9 from each ward. A quota was fixed so that half of the respondents were male, one third were aged 35-44 years, one third were aged 45-54 and one third 55-64. In general, the pilot was successful. In two weeks 80 respondents were surveyed face-to-face according to quota, of whom 47 were "primaries". In

total 384 houses were approached, 184 of whom were out of quota, 80 refused an interview, 49 were unavailable, 9 were empty and 2 were too sick or disabled. After analysis of the results minor changes were made to the sampling rules, the logic and wording of the questionnaire, and to the system of coding. Interviewers were warned of inconsistencies and of areas where probing the initial responses might lead to greater clarity. There was an indication that memories were more selective about recall with distance in time, since jobs in the distant past were generally claimed to have been of longer duration. This may be a real finding but it may also show that retrospective histories tend to be more detailed about the recent past. The biggest problem was the collection of information about live-in partners. The questionnaire did not distinguish cohabitation from marriage and separation from divorce, in order to avoid the common bias and loss of data stemming from concentration on marriage (Tuma 1994). However, several respondents were uneasy about discussing past relationships when their current partner was in earshot. As a result, the question was simplified even further.

The method of sampling in the pilot was successful, picking up "status zero" respondents and travellers who had never been employed for more than a few weeks at a time. These may be the people left out of other sampling designs which focus on those in work and more stable conditions. In general, the findings reflect the conclusions drawn from the secondary data (Gorard 1997a). All of the respondents defined themselves as "white", nearly all were married (or had been married), few had more than one previous relationship, and most had always lived in South Wales. Over three quarters had no full-time education since compulsory schooling, and nearly half had no substantive training either. There were no graduates. Most jobs involved no training longer than a few days. There was a clear cohort effect in terms of the type of school attended, which led directly to differences in terms of qualifications, truancy, continuation to post-compulsory education, and later courses of adult education. This initial pattern also affected the types of institutions used for all subsequent education and training, and the reasons reported for studying. The pilot confirmed the existence of locally mediated trajectories for post-compulsory education and training as well as cohort and period effects. A simple attempt at scaling the results suggested that at least one way of dividing the patterning in the data is to identify a dimension which distinguishes older, Welsh-speaking, local born, non-movers from those who move more regularly and are from outside the region. This

second group generally has more qualifications and fewer truants. A second dimension separated those who take later voluntary training courses from those who do not. These people tended to have higher occupational status and social class, to work for larger companies, supervise others, and gain promotion as a result of training.

A simple typology of learning histories was created based upon the sequences of individuals participation in education and training. This divided the respondents into six groups (Gorard et al. 1996) ranging from those who reported no formal education or training since reaching school-leaving age to those who had moved in and out of formal learning situations throughout their working lives. There were systematic differences between these groups in terms of characteristics not used in their definition - for example, the lifetime learners more frequently moved to the research area from outside South Wales - and these reinforced the validity of the procedure used. This method of categorising learning histories became the basis of that used in the main study.

Coding

The answers to the questionnaire provided a different number of variables per case, depending on the number of episodes in the work and learning histories. On average, each respondent produced several hundred simple variables. Some of these were in text format, requiring recoding before analysis. For example, the description of employer's industry was converted into a Standard Industrial Classification code, and descriptions of occupations were converted into both Occupational Class and Social Class codes. Some variables required the creation of new coding systems, such as that for long-term hobbies requiring self-study. For most other questions the coding scheme was implicit in the list of prompts provided with them. In some cases respondents had replied to closed questions with "other" and then specified an answer that was clearly the same as one of the prompts. These were recoded as necessary. The designed scales for some variables were collapsed to eliminate sparsely populated cells (Lee et al. 1989), and minimise the number of structural and random zeroes in further analysis (Norusis 1994), including for example those for first language, ethnicity and family religion.

Some completely new variables were created at the coding stage (see Appendix). Sometimes these involved a simple calculation or conversion, such as changing the years in which events happened into the age of the respondent at the time, or providing a count of the variable-length-per-case data, such as the number of children. Some more complex variables were also derived, such as the proportion of time spent in all formal learning and some were combined to create proxy measures for more esoteric concepts (cf. Davies 1994). A typology of trajectories was created, as in Banks et al. (1992) but the learning trajectories identified here were more complex covering a time span of 50, rather than three, years (forthcoming paper).

3. ANALYSIS

Several forms of analyses have been carried out with the results from the main study, including the creation of descriptive statistics for each individual variable, comparison of the values of combinations of two variables within each case, consideration of interaction effects, reduction of groups of two or more variables to underlying explanatory variables and the creation of a range of hypothetical multivariate models with dependent variables. These analyses have been carried out with different versions of the data file, including treating each individual record as a case, treating each parent and child combination as a case, using each reported work episode as a case and using each reported training episode as a case. One of the disadvantages of working with data at a high level of abstraction was that many variables had few valid responses, since they formed part of a variable-length history, such as the occupational class code for the *n*th job since leaving full-time continuous education which only existed for those with at least *n* jobs. One of the disadvantages of working at a lower level was that the individual work and training episodes were not related to the family history and other life events of the individual. The solution was therefore a compromise between the ideal and the practical. It is common practice for samples to be separated by gender for analysis, since their patterns of participation in education and training have traditionally been so different (Dolton et al. 1994), perhaps partly for cultural reasons or simply due to different levels of access to motorised transport. Although it was considered for this survey, such a separation of cases could also be argued for each research site, for each time period, or even for marital status (cf. Greenhalgh and Stewart 1987) leading to rapidly dwindling numbers of cases per cell. Since changes over time in the differences between genders are

themselves an area of interest for this study, most of the analysis was completed for the whole sample, and subsets were only used where indicated by the primary findings.

The majority of the data, which was in ordinal and categorical form, was summarised as frequencies (Frude 1993), and converted to graphical form where appropriate. These frequencies characterised the sample. They were compared to the known population parameters (Gorard 1997a), used to detect outliers, and used in the consideration of corrective weights for each strata where necessary. Each of the simple variables was converted into percentages for standardisation (Reynolds 1977), cross-tabulated with every other (Gilbert 1993), and where a relationship of interest appeared, tested for significance by means of chi-squared (Siegel 1956, Reynolds 1977). Those items measured on an interval scale were summarised as mean scores with standard deviations while the relationships between them were estimated by correlation coefficients (Gorsuch 1972). The mean of each interval variable was calculated for each sub-group of respondents defined by each categorical variable, and the significance of potential differences between them assessed by t-tests or one-way analysis of variance, as appropriate (Levine 1991). Some of the emergent two-way relationships - older people having more children - would have been clearly expected under almost any reasonable hypothesis (Wisdom 1987) and these simply reinforced the reliability of the survey. Others were research findings of some importance.

More complex associations were tested by simple means such as the Kendall coefficient of concordance measuring an association between three or more variables (Siegel 1956) and three-way crosstabulation of data. One method used to uncover a hidden structure in people's behaviour was to draw a map of the distances (or proximities) between the values of variables, measured as points - multidimensional scaling (MDS). Thus the data were converted to correlations and thereby to measures of distance, and drawn on a map of a specified number of dimensions. Several solutions were attempted, each with a different number of dimensions. The quality of the solutions were assessed by their stress, a measure of badness of fit, by their usefulness in explanation (Kruskal and Wish 1978) and by their squared correlations (Schiffman et al. 1981). A plot of the number of dimensions in each solution against their stress and correlations revealed an "elbow" pointing to the maximum number of dimensions discernible in the

solution (Schiffman et al. 1981). In practice a map of more than three dimensions was unhelpful, as it was difficult to interpret and unstable (Kruskal and Wish 1978). Although plotted against x,y axes, several other candidate lines in space were used to separate clusters of items that were describably different to each other. MDS is a little like factor analysis in that it elucidates factors or hypothetical dimensions underlying the real data. This can simplify the picture and reduce the number of variables for analysis and perhaps the problems of multicollinearity before regression.

Factor Analysis

Principal Components Analysis (PCA) of the many interval variables in the data set followed for those with sufficient responses, using a similar method to that described in Gorard (1996). This reduced the many years, ages, dates and times, among other variables, to a few underlying hypothetical factors for use in describing normative lifestyles, thereby identifying idiosyncratic cases for further investigation, via in-depth interview for example. The variables used in this exploratory PCA were either clearly measured on an interval scale, such as the year of birth, or were binary or dichotomised variables, such whether married or not, all of which can be used for factor analysis despite having an undefined metric (Kim and Mueller 1978$_b$). It was not necessary for the data to be normally distributed (Kim and Mueller 1978$_b$), although they were approximately, nor to show linear regression between all pairs of variables (Comrey 1973), although most of them did. There were sufficient cases for factor analysis since Comrey (1973) recommended at least 300 cases and there were over 1,100, and Child (1970) suggested the need for at least five times as many cases as variables and fewer than 100 variables were used. It was not, however, possible to rule out the possibility of bias even in well-loaded factors since the variables concerned may fail to represent some important aspect of the emergent underlying pattern. Some information was lost in PCA, but most of what was lost are the idiosyncratic, erratic, or irrelevant components of the variance in responses (Marradi 1981), and since the use of multiple tests tends to lead to spurious results, factor analysis can combat this by reducing the number of potential tests (Stevens 1992).

The first step was to decide which of the available measures to include in the analysis. A matrix of correlation coefficients was produced, and examined for an intuitive feel for the data. Other types of correlation or covariance matrices could

have been calculated (Siegel 1956), but the standard Pearson r is the most common, and the most appropriate to these data (Comrey 1973, Kim and Mueller 1978_a). One purpose of the analysis was to find a smaller number of satisfactory substitute variables, so substantial correlations between variables were required (Comrey 1973). A criterion was needed to determine which correlations were large enough to be worth further investigation, and which to exclude from the reduced correlation matrix (Marradi 1981). To exclude the clearly unrelated variables can lead to a neater solution, but it can also damage the result. On the other hand, to include them would be to court unnecessary complication, and was contra-indicated by the principle of "garbage-in garbage-out" (Kim and Mueller 1987_a). With such a large sample several minor factors appeared as statistically significant, without contributing greatly to the overall covariance structure (Kim and Mueller 1987_a). All researchers have to make a discretionary judgement at this point and the method chosen here was to select a figure for the size of common variance that could be useful for explanatory purposes, such as 10%, calculate the size of the correlation that represents it, such as |0.3162|, and exclude all variables that have no correlation of at least that size. The selection and discarding of measures at this stage is too important to be left to a computer, and is of necessity an iterative process (Marradi 1981), so a final decision was not made until all the data were collected, and it was discovered how many of the variables to drop as too "complex" for their explanatory power (Comrey 1973). Further uses of the correlation matrix were to confirm that it was not an identity matrix and to calculate and check Kaiser's measure of sampling adequacy.

The second step was to extract the unrotated factors underlying the variables. The Principal Components method was used as it is good for extracting the maximum variance (Gorsuch 1972), more appropriate for reducing the number of variables, especially where the first few factors account for a large part of the variance, and more useful as a prelude to further analysis. PCA is also "*straightforward in the sense that no distributional assumptions need to be made about the observed variates*" (Maxwell 1977 p.42). Anyway, the same solution is often rotated to from other starting points based on other methods of extraction (Gorsuch 1972), and whichever method is used initially, the subsequent rotation of the factors leads to similar results (Comrey 1973). While the number of factors were kept to a minimum (Cureton and D'Agostino 1983), several needed to be extracted to create a good enough structure for rotation (Comrey 1973).

Kaiser's criterion was used to define a cut-off point for the number of factors extracted.

The next step was to rotate the factors to a solution, which although mathematically equivalent to the original, was more useful for explanatory purposes. The main choice was between an oblique rotation, allowing the factors to overlap, or an orthogonal rotation, forcing the factors in the solution to be unrelated (Child 1970). An oblique rotation would be necessary if higher order factors were to be derived, but the gain in generalisation from second order factor analysis is at the cost of less accurate results (Gorsuch 1972). As with many decisions in the analysis, and perfectly properly, a mathematical decision to use orthogonal rotation was taken for non-mathematical reasons (Comrey 1973). Several useful criteria for deciding on a method of rotation are: the ease of interpreting the results; the number of zero loadings; the speed of convergence; the restriction of each variable to only one factor, and the replication of factors across split halves of the cases (Gorsuch 1972). The intention here was to produce a small number of clearly unrelated factors underlying the life histories, with clearly differentiated loadings. This does not mean that an oblique rotation would not produce meaningful results, merely that they may not be so clear cut. A set of uncorrelated factors can also be seen as more parsimonious (Cureton and D'Agostino 1983).

The final step was to interpret the factors produced by the rotation. This involved both naming and explaining the factors, each of which was then treated as a hypothesis to be tested further. It is important to decide of each factor - is it valuable, or is it an artefact? This question can only be answered by further analysis and comparison with other data sources. Each variable in the analysis has a final loading on each factor. The loadings are the correlations between variables and the factors, so the square of a loading can be seen as the amount of variance common to both. A reasonable cut-off point would be 0.3, with loadings below that figure being ignored in explanation of a factor (Child 1970, Comrey 1973). This study uses the more precise, and more stringent, figure of |0.3162|, which is the square root of 0.1, so that variables used in the final model have at least 10% common variance with the factors to which they contribute. However, the threshold for a loading cannot be defined rigidly in advance, as it really depends on where a sharp drop in values is noticed (Marradi 1981). Since measurement of

the variables was unlikely to be wholly reliable, they do not correlate perfectly with any factor, but as a rule, if the loading of a variable is equal to the square root of its reliability, it is a "pure factor" measure, i.e. the variable and the factor are identical, and it is variables of this type that are most useful in shaping the factor "space" (Comrey 1973).

Once a satisfactory set of factors were derived from the data, they were subject to further analysis. The factor scores per case were calculated and saved using the most common method of calculating factor scores after PCA (Norusis 1994). These scores are defined as the sum of the case score on each variable, multiplied by the loading for that variable on the factor, for all relevant variables (Jackson and Borgatta 1981). Unfortunately, this procedure reduced the number of cases, since only cases with valid responses to all of the relevant variables were given a score for the factor. One-way analysis of variance, and t-tests were used to test for relationships between the factor scores and the nominal variables.

All of the above techniques either do not have a dependent variable, or the dependent variable can be described as the observations themselves, for example the observed cell counts. Further analysis involving predicting or explaining the variance of one measure in terms of others requires an initial consideration of causal models.

Causation

In research, as in life, an easy assumption is sometimes made about the direction of causation that does not really stand up to analysis. This assumption is that one event can only be considered a cause of another if it occurs first, therefore if two variables are related then their temporal sequence defines which is the cause. For example, an analysis by Dolton et al. (1994) of data from the longitudinal Youth Cohort Study explained the labour market position of each participant in terms of their position in previous sweeps. Similarly, Gershuny and Marsh (1994) explained each participants' current employment position, as described by employment status, sector and occupational level, in terms of two main determinants, both preceding the current employment position - their initial characteristics and the accumulation of previous employment. They therefore adopted a "recursive determination model", which was based on the causal chain approach advocated by Blau and Duncan (1967). Technical literature generally

suggests methods for analysis of event histories, such as Proportional Hazard modelling (Norusis 1994), which try and explain occurrence variables in terms of prior events. In fact, a majority of studies use these unidirectional recursive models, such that the predictors or independent variables are themselves unaffected by the outcomes or dependent variables (Berry 1984). The approach was summed up in one study thus, "*what we do now becomes what we are, and what we are in part determines what we do next*" (Gershuny and Marsh 1994, p. 69). In their analysis of the determinants of unemployment, variables were entered into the model in the order that they occurred historically, from parents' occupational class through initial education to the work details. The "effect" of the earlier episodes was assumed to be present throughout the analysis but was found to diminish over time. In this way, the past is seen as affecting the present while both can affect the future, but the future cannot affect the present and the present cannot affect the past. To a large extent the simple view of causation denies the possibility of choice, but rather sees reproductive forces pushing people into pre-determined educational and career routes.

However, in many respects the assumption of unidirectional causation is unrealistic (Berry 1984). Causality is merely assumed to be time-determined (Hume 1962) and like induction remains a perpetual "skeleton in the cupboard" of philosophy. Also, the relationships between data which are seemingly in a temporal sequence are often reciprocal (Hagenaars 1990). Day always precedes night and so could explain 100% of the variance in a regression model, but that would not make it the cause. Many educational decisions may be made in the expectation of a fairly remote future outcome, for example a decision to stay on at school at age 16 in order to become a barrister. Even as early a stage as picking courses of study for GCSE can be dependent on final career intentions (Roker 1991). Rational choices can allow people to jump towards attractive options rather than being "*pushed from behind*" (Gambetta 1987). In addition, the direction of the arrow of causation is not at all clear even in well-established links between variables. For example, does a higher family income lead to a better education for the children, or can stress laid on future educational plans also lead to a need for a higher level of income? Analyses using only a single equation model, such that A is predicted by B and C separately assume no relationship between the predictors (Berry 1984). Regression models are however, becoming more complex with multi-equation rather than single equation models, which

allow the predictors to influence each other, so that although A may be caused by B, both A and B may be caused by C. However, in theory at least, if two variables can be reciprocally related or even if their error terms are related, then a non-recursive model must be used, but this is seldom seen in social science outside econometrics (Berry 1984). One reason for this is that as it is not possible to deduce the direction of causation from a simple association, some non-recursive models cannot be meaningfully described. Perhaps this is why research generally explores associations between family background, education, occupation and earnings as though these events were sequential and causally dependent, whereas in reality, people may leave and enter education and work a number of times in their lives and even carry out both at the same time for example (Mare and Winship 1985).

A more complex solution is to construct a model that involves both causation and other competing explanations of an non-determinist nature, such as intentionality through personal choice. Gambetta (1987) describes educational decisions as a product of what is available to the individual, what the individual wants and, indirectly, the social conditions which shape the individuals intentions. However, explanations such as this are unparsimonious on two counts. Firstly, within the causal model, if a cause can be either direct or indirect, an infinite number of possible intermediate steps can always be created between the observed direct "cause" and the hypothesised indirect one (Blalock 1964). If either explanation fits the data, the simplest solution is the best and the simplest solution cannot be both solutions at the same time. Similarly, the problem with causation is not that there are events that it cannot explain, but that it is impossible to measure. Therefore there is no value in mixing it up with a model of intention which is also perfectly capable of explaining decisions by itself but which is also not open to observation by social scientists. Given that there is no way of deciding between them empirically, either causation or intention can be adopted (it makes little practical difference which at this stage). There is no empirical reason to work with both at the same time.

Causation is therefore a difficult concept for use in social science for a variety of reasons. It is difficult to define, requiring some form of agency or productive force to be meaningful, although this force is never observed and merely induced from patterns of ordered events (Blalock 1964). Isolation of variables is very

difficult in the non-experimental situations common to social science research and, as seen above, can lead to the lack of a clear temporal sequence. However in some ways causal models are very simple, requiring what may be unpalatable assumptions for some researchers (Blalock 1964), since in actual research causal inferences are often drawn with no clear rationale. For some, the rhetorical trick of defining one variable as the dependent one in a formula is sufficient to "prove" that it is an effect as long as a satisfactory regression line can be fitted. However, mathematical dependency is a matter of convenience only and any function of the form $Y = f(X)$ can be rewritten as $X = f'(Y)$. What would such a recasting do to the assumptions underlying the findings of many past studies?

Given the difficulty of identifying causes, perhaps the best that can be hoped for is to identify only weak causes or "determinants". These could be the producers of the observed effects, or they could be simply the indicators, or sign posts, of a future outcome. There is no way of distinguishing between these possibilities, therefore there is no point in this study trying to do so. In this way, the model of determinants emerging from this study can be seen as either causatory or indicative of personal choice and each will be considered later in relation to the other data sources. Whichever description is used, the determinants must be able to predict the outcomes accurately. This alone will be the basis for their claims to be taken as valid explanations.

Temporal sequences

It is impossible to analyse the life histories collected by this survey without a consideration of time. Variables, such as the level of qualifications attained during any learning episode or the size of the company in a work episode were compared within one individual case to describe patterns over one lifetime. Changes in the frequencies of these patterns across individuals and regions were also examined over the period of time covered by the study. The final state of each history (e.g. the final occupational class) was analysed as a simple variable, as with a standard cross-sectional survey, but it was also necessary to include time in the calculations in order to look for possible period, age and cohort effects. This was attempted by a series of "cross-sectional" simple analyses at different time periods (such as after the birth of a child, or upon leaving full-time initial education) as well as through a more general model with time-dependent variables (Allison 1984).

During the lives being reported in this study, social and economic changes were taking place which could have had a direct bearing on them. Individuals working and training in the "booming" 60s may have very different experiences despite having similar background characteristics to those of the "Thatcher" 80s (Gershuny and Marsh 1994). For this reason the sample was divided into age cohorts. Standard regression was not used to analyse changes over time, since in any standard cohort table, with rows for each age group and columns for each cross-sectional date, at least three effects are confounded (Glenn 1977) and it was important to distinguish between them (Blau and Duncan 1967). In an extreme example, where the table shows pure age effects each column will have identical values. A period effect would show only variation by column and no variation by rows. A birth cohort effect would show variation by both rows and columns, but would show no variation by the upper left to lower right diagonals. Of course, each effect could also be described as a compound of the other two, but this would be less parsimonious and was therefore not attempted in the first instance. The effects of time are always complex, with the age sensitivity of some variables like length of study in initial schooling leading to the appearance of age sensitivity in the labour market (Mare and Winship 1985). It is not easy to untangle these effects but as long as the researcher has a knowledge of the effects of ageing and the history relevant to the period of study, as in this study, the analysis requires nothing more complex than a chi-square test on the standard tables (Glenn 1977). However, like regression and many other common treatments, analysis of variance is unwise in event history analysis because of the problems of right censoring - participants not yet arriving at the event (Tuma 1994) - and the variation over time in explanatory variables (Allison 1984).

In the analysis, time could have been treated as a discrete variable, which is perhaps more natural but required an accuracy in the measuring and timing of events that would be uncommon in social science (Allison 1984). Thus, since errors must be assumed in the dates and times appearing in the retrospective histories gathered by this study and since there are also several ties or simultaneous events, time is treated as a discrete variable with all events measured only by their year of occurrence. The results of both approaches are usually similar anyway (Allison 1984). Such a broad brush approach was used by Gershuny and Marsh (1994) who grouped their sample into birth cohorts of 10

years range, and ignored small-scale variations in measurement as being symptomatic of random error. Their plots of unemployment are therefore smoothed, and because of the simplification they managed to detect both cohort and period effects in the data. As explained above, the analysis here involved separate treatment of each type of repeated event, treating each training episode as a case in itself (Allison 1984) as well as a more complex treatment of multiple types of events, such as child birth or educational "failure", where the occurrence of one event was allowed to affect another (Allison 1984). The values of variables at different times were made into variables themselves, thus the highest qualification at age 27 and at age 28 might be two variables for example. This reduced the flexibility of the system and made it easier to work with. In this way there was no necessity to use a non-recursive model, even though reciprocal causation was permitted and some dependent variables were also independent variables for the same model (Blalock 1964).

Standard regression

Standard multiple regression calculates the probability of an event or outcome as a function of:

$B_0 + B_1X_1 + B_2X_2 + ...$ where X_1, X_2 are independent variables and B_0, B_1 and B_2 are calculated from the data using the method of ordinary least squares to minimise the deviation between the observations and the predictions. Multiple regression is similar to a series of individual analyses with the advantages of allowing interaction effects between predictors, reducing the amount of unexplained variance (Pedhazur 1982) and avoiding the problems of repetition of tests originally designed for one-off use (Gorard 1997b). An example of this is in the Mincer equation for earnings, where the natural logarithm of a person's current earnings are seen as a quantifiable function of their past training, personal characteristics, and other factors including local market conditions (Dolton et al. 1994). However this approach is not appropriate for non-recursive models (Berry 1984), which is one reason why consideration of the causal model must precede analysis. Standard regression also does not work well with polychotomous categorical independent variables (Hagenaars 1990), such as the typology of learning trajectories proposed by this study. The regression model assumes a linear relationships between interval variables taken from a random sample, with no interaction effects and strict assumptions about the distribution of the error term (Maxwell 1977). Achen (1982) agrees that standard regression

should not be used where the parameters are not linearly related to the dependent variable, as they seldom are. There is some dispute about the validity of overcoming this by converting categorical variables to a series of dichotomous independent variables or converting non-linear relationships via the use of logarithms, as is common practice in some fields. Some writers have argued that since it is assumed in regression that the dependent variable is normally distributed (Lee et al. 1989), and dummy variables cannot have such a distribution (Hagenaars 1990) simply converting a categorical variable into a set of dummies does not help. It also adds to the measurement error (Blalock 1964). Designs with categorical independent variables have traditionally used a series of analyses of variance, but this has been shown to be special case of multiple regression (Pedhazur 1982) and the pressure to collapse and dichotomise often variables remains (e.g. Gershuny and Marsh 1994), so that regression is considered by some as equally appropriate for both continuous and dichotomous data.

Statistical textbooks describe ideal procedures to be followed, but several studies of actual behaviour have observed different common practices among researchers (Gephart 1988). In any real research project involving multiple regression, some of the assumptions underlying it are likely to be violated (Berry and Feldman 1985). This, in itself is not fatal to the validity of the findings, but it is important to be aware of the assumptions that are violated and the impact of this on the analysis. In many cases, even where the regression is deeply flawed it is only the intercept that is affected and the derived coefficients may still be used with care, e.g. where the means of the error terms are not zero. Of the other six assumptions listed by Berry and Feldman (1985), the error term for each set of values for the independent variables should be normally distributed, but this is ignored for large samples due to central limit theory. Regression is robust against violations of normality. No independent variable should be allowed to be perfectly linearly related to any other, but this possibility is so rare that it can be ignored in practice (Berry and Feldman 1985), although imperfect linearity (multicollinearity) is more common. All variables should be measured without error, so that in theory a low value of explained variance in the solution (R squared) shows that theoretically relevant variables have been left out of the model. In practice, of course, there are measurement and other errors such as incorrect specification of the form of the equation meaning that low values of R

squared can be accepted at face value. In general the relevance of variables in the solution is theory-driven, although those with coefficients indistinguishable from zero can be dropped (Berry and Feldman 1985). The independent variables should be linearly related and additive and these assumptions are also best determined by theory. Other assumptions concerning the nature and relationships of the error terms were assessed by plotting the residuals which should have an even symmetric appearance. If not, a formal test for heteroskedasticity was conducted.

Despite the caveats about the appropriateness of various methods, there is always an element of trial and error, while the knowledge of whether analysis has worked is almost intuitive. A good model fits the data. It is easy to interpret and parsimonious in its assumptions (Norusis 1994) and it triangulates well with other data sources. Factors such as these along with other practical considerations such as the ease of variable acquisition can be used to decide between competing alternatives (Norusis 1994). A method is satisfactory if it reveals verifiable theoretical understanding in a field and good data analysis mainly stems from skills other than mathematical abstractions. Progress is not ensured by methodological advance, while on the other hand many good pieces of work are based upon analysis that a mathematician would mock (Achen 1982). With this in mind, regression was attempted with a subset of the data from this survey. Analyses included an estimate of the predictors for the number of years spent by a respondent in full-time continuous education, for the proportion of one's lifetime spent in learning episodes and for the amount of work-based training available in each employer/grade/industry/task combination. In this way the research addresses questions such as what determines staying on in education, how can one recognise a lifelong learner and what types of training schemes have been available in which types of jobs?

Logistic regression

One thing that regression cannot be used for is to model a categorical dependent variable (Achen 1982). However, where the dependent variable is, or can be converted to, a dichotomy such as whether an individual was unemployed or not in the preceding year (Davies 1994, Gershuny and Marsh 1994) a range of regression-type techniques are available. Several recent studies have used the probit model (Mare and Winship 1985, Main and Shelly 1990, Whitfield and

Bourlakis 1991) with simplified and categorised variables that would not be appropriate for linear regression, while others recommend the use of logistic regression for the analysis of time-dependent variables (Allison 1984). Both techniques are similar, using the predictors to compute a score on an underlying latent variable. If this score is above a specified critical value the dependent variable is set to one category, else the other (Whitfield and Bourlakis 1991). Logistic regression requires fewer assumptions than many tests (in particular the variables need not be linearly related) and is more valid than standard regression for categorical dependent variables (Norusis 1994). Logistic regression was used here to create a series of models, such as the determinants of staying on at school after compulsory school-leaving age, taking terminal qualifications or not, returning to education in later life or not, and receiving substantive work-based training or not. The predictors used included family background, past educational experience and current status, entered into the model in several ways including in the order in which they occurred, so that a sequence of forward causation could be assessed.

Logistic regression sets the probability of an event as:
$1 / [1 + e^{-(B_0 + B_1X_1 + B_2X_2....)}]$
Thus the log of [the probability of an event/probability of no event] = $B_0 + B_1X_1 + B_2X_2$ ... known as the log-odds where the B coefficients are calculated from the data iteratively since this is not a linear model like standard regression (Norusis 1994). Both weighted-least-squares and the method of maximum-likelihood can be used in the calculation and they give similar results in most cases, but maximum-likelihood was used here as it can be applied to much larger tables (Lee et al. 1989) and it maximised the probability of observing the data obtained (Allison 1984). The R statistic obtained shows the contribution of each variable to the event. A positive value of R increases the chance of the event, and small values show that the variable is of little relevance or that may be related to other predictors (Norusis 1994). In this way redundant and clustering variables were eliminated at this stage (Lehtonen and Pahkinen 1995). For large samples the test that any coefficient is actually zero used the Wald statistic which has a chi-squared distribution, with degrees of freedom equal to the number of categories in the variable minus 1. If the probability of the Wald statistic was less than 5%, the coefficient was assumed to be non-zero and so used in the model as a predictor. However, this test is less accurate when the coefficient is very large, so

another method was to try the model both with and without the coefficient and look at the change in log-likelihood. Twice the difference between the log-likelihoods of two competing models has a chi-squared distribution with degrees of freedom equal to the difference between the number variables in each model (Allison 1984). Where one model was significantly worse it was rejected, but where both models fitted the data equally well, the simplest was selected in accordance with the principles emerging from the discussion on causation.

Loglinear models

Perhaps one reason why standard regression has been used in some past studies of education is that education in the USA has a yearly basis, making the number of years in education a simple dependent variable to model (Gambetta 1987). In the UK the situation is different, but on the other hand choices are seldom binary in nature and to constrain them as so for the purpose of analysis may increase measurement error. Logit loglinear methods, on the other hand, allow decisions to be represented as discrete but non-binary choice sets. Loglinear methods have been particularly recommended for social science where so many of the variables are categorical in nature (Gilbert 1993). Whenever an analyst is faced with three or more variables and their possible interaction effects, a test like chi-square would be overused (Stevens 1992). Rather, an iterative application of the hierarchical loglinear method is appropriate, using the log likelihood ratio statistic G square to assess the goodness of fit of each loglinear model. This prevents needless repetition of the chi-squared test, allows interaction effects, and produces an estimate of the size of each effect (Norusis 1994). The models are hierarchical since all terms in the model imply that all lower order terms must be included, so that the interaction AB implies the main effects A + B individually.

In general, loglinear analysis looks at the patterns in all possible partial tables in a multi-dimensional cross-tabulation of data. Using the table of observed values, the table of expected values is constructed iteratively such that it fits all known marginal totals (Gilbert 1981). Three dimensional tables (tabulating three variables) have a total of 19 possible model to explain them, ranging from the grand mean, where each cell is expected to be identical, through the simple effect of each variable in isolation, to a model in which every possible interaction is accounted for, e.g. A + B + C + AB + AC + BC + ABC. For each of these possibilities the log likelihood ratio statistic, G square, is calculated and looked up

in a chi square table, with degrees of freedom equal to the number of cells minus the number of fitted marginal totals. A well-fitting model will have a high probability, but to test how good the fit actually is, the solution can be compared to the grand mean to produce a coefficient of multiple determination, where 0.9 explains 90% of the variance for example. There are several approaches to finding the best fit. In forward selection the grand mean model is examined first, followed by successively more complex arrangements. The simplest model which has been found to be likely, i.e. of more than 5% probability, is selected but then may be modified by the addition of some of the interactions from more complex solutions (Gilbert 1981).

The dependent variable can be the number of cases in a cell of a crosstabulation of independent variables. The log of this cell value is the average of the logs of all frequencies (mu) plus the change (lambda) from mu for each row/column combination (main effect) plus the lambda for interaction, which is the log of the cell count minus mu minus the effect for row minus the effect for column. All of this is estimated from the observed data, as the lambdas must total to zero across all categories (Norusis 1994). For this reason, binary values are treated as opposites. When the model includes all the main effects and all the interactions, it reproduces the observed counts exactly, and is said to be saturated. The estimate of each parameter divided by its Standard Error (the Z value) can be used to test the null hypothesis that each lambda is zero. If modulus Z is greater than 1.96 then lambda is significant at the 5% level, and a 95% confidence interval can be calculated for the value of that parameter. Parameters with small values were therefore excluded from the saturated model (Norusis 1994), leading to a simpler, more practical solution.

Where the cell frequency was zero, cells were either collapsed or 0.5 was added to the empty cell (Hagenaars 1990). Each expected value was calculated iteratively using the standard convergence criterion of the difference between successive estimates. Chi-square is used to test the difference between the observed and expected scores and estimate the goodness of fit of the model, with degrees of freedom equal to the number of cells minus the number of parameters (including the grand mean) plus 1. If the chi-square probability is small, the model is rejected. Another way of testing for goodness of fit is to standardise the residuals, the difference between observed and expected, by dividing each by the

standard deviation of the observed values. If the modulus residuals are greater than 1.96, the model is rejected, and this has the advantage that which residuals are too large may give a clue to the problem, often with the use of a plot.

The method can also be used to predict the value of a nominal dependent variable having more than two categories, by applying the multinomial model of logistic regression, and finding the simplest logit model that fits well (Gilbert 1993), i.e. where there is a response variable with several predictor variables including continuous covariates (Greenhalgh and Stewart 1987, Norusis 1994). These logit loglinear models assess the effect of all variables considered simultaneously, each net of the effect of the other, unlike the standard linear distribution (Gambetta 1987). The response variable can be binary, polychotomous or continuous (Lehtonen and Pahkinen 1995) and the typology of training trajectories was used here as a polychotomous categorical dependent variable, summarising the pattern in the observations as a predictive function with coefficients for all of the significant parameters. Unlike the standard loglinear method, which counts the occurrence of the dependent variable for each combination of independent variables, the logit models the ratios of the counts. It has fewer parameters since it only provides estimates for parameters involving the dependent variable. The size of the coefficients in the model may not be directly proportional to their effects, but these odds ratios show whether an "effect" is present and whether it is positive or negative as well as providing an idea of its strength (cf. Gambetta 1987). In this way the social determinants of adult education and training trajectories can be identified.

CONCLUSION

The design of the remaining components of this study of patterns of participation in adult learning, such as the in-depth interviews and analysis of changes in local training provision are the subject of a later paper. In summary, the survey component focuses on the structural and social determinants of participation for individuals and their changes over time. The background constraints, the "training market", and above all the attitudes of non-participants may appear as unexplained variance in the analysis and are themselves the focus of the later components of the study. Confirmation of the value of the emerging patterns will come from split-half analysis, predictions to be tested by planned future waves of the survey and triangulation with other data sources. In order to give the findings

some substance, the marginal effect of each variable in the logit model was calculated for a hypothetical base individual (cf. Greenhalgh and Stewart 1987, Whitfield and Bourlakis 1991) to give a useful estimate of the effect of each variable on participation in lifelong learning. Such a picture can make the policy-relevance of the findings startlingly clear.

ACKNOWLEDGEMENTS

The ESRC (Grant Number L123251041), Gwent, Mid Glamorgan and West Wales Training Enterprise Councils for funding. The families and training providers for participating.

Address for correspondence
Stephen Gorard
School of Education
21 Senghennydd Road
Cardiff CF2 4YG
01222-874000

REFERENCES

Achen C. (1982) **Interpreting and Using Regression**, London: Sage

Allison P. (1984) **Event History Analysis: Regression for longitudinal event data**, London: Sage

Banks M., Bates I, Breakwell G., Bynner J. and Emler N. (1992) **Careers and Identities**, Milton Keynes: Open University Press

Berry W. (1984) **Nonrecursive Causal Models**, London: Sage

Berry W. and Feldman S. (1985) **Multiple Regression in Practice**, London: Sage

Blalock H. (1964) **Causal Inferences in Nonexperimental Research**, Chapel Hill: University of North Carolina Press

Blau P. and Duncan O. (1967) **The American Occupational Structure**, London: John Wiley

Child D. (1970) **The Essentials of Factor Analysis**, London: Holt, Rínehart and Winston

Comrey A. (1973) **A First Course on Factor Analysis,** London: Academic Press

Crouchley R. (1987) **Longitudinal Data Analysis. Surrey Conferences on Sociological Theory and Method 4**, Aldershot: Avebury

Cureton E. and D'Agostino R. (1983) **Factor Analysis: An applied approach**, London: Lawrence Erlbaum

Dale A. and Davies R. (1994) **Analyzing Social and Political Change: A casebook of methods**, London: Sage

Davies R. (1994) From Cross-Sectional to Longitudinal Analysis, p. 20 in Dale A. and Davies R. (Eds.) **Analyzing Social and Political Change: A casebook of methods**, London: Sage

Deloitte, Haskins and Sells (1989) **Training in Britain. A study of funding, activity and attitudes. Employers attitudes**, London: HMSO

Dolton P., Makepeace G. and Treble J. (1994) Measuring the effects of training in the Youth Cohort Study, in McNabb R. and Whitfield K. (Eds.) **The Market for Training**, Aldershot: Avebury, p. 195

FEU (1993) **Paying their way: the experiences of adult learners in vocational education and training in FE Colleges**, London: Further Education Unit

Frude N. (1993) **A Guide to SPSS**, London: Macmillan

Gambetta D. (1987) **Were they pushed or did they jump? Individual decision mechanisms in education**, London: Cambridge University Press

Gephart R. (1988) **Ethnostatistics: Qualitative foundations for quantitative research**, London: Sage

Gershuny J. and Marsh C. (1994) Unemployment in Work Histories, p.66, in Gallie D., Marsh C. and Vogler C. (Eds.) **Social Change and the Experience of Unemployment**, Oxford: Oxford University Press

Gilbert N. (1981) **Modelling Society: An introduction to loglinear analysis for social researchers**, London: George Allen and Unwin

Gilbert N. (1993) **Analysing Tabular Data: Loglinear and logistic models for social researchers**, London: UCL Press

Glenn N. (1977) **Cohort Analysis**, London: Sage

Gorard S., Rees G., Fevre R., and Furlong J. (1996) **The social determinants of adult participation in education and training: a framework for analysis**, paper presented to British Educational Research Association Annual Conference, Lancaster - also available on Educati*on-line*

Gorard S. (1996) Three Steps to "Heaven"?: The family and school choice, in **Educational Review**, 48, 3, pp. 237-252

Gorard S. (1997$_a$) **Patterns of participation in adult education and training. Working paper 1: The region of study**, Cardiff: School of Education, ISBN 1 872330 02 9

Gorard S. (1997$_b$) A choice of methods: the methodology of choice, in **Research in Education**, 57, pp. 73-88

Greenhalgh C. and Mavrotas G. (1994) Workforce Training in the Thatcher Era - Market forces and market failures, in McNabb R. and Whitfield K. (Eds.) **The Market for Training**, Aldershot: Avebury, p. 19

Greenhalgh C. and Stewart M. (1987) The effects and determinants of training, in **Oxford Bulletin of Economics and Statistics**, 49, 2, pp. 171-190

Hagenaars J. (1990) **Categorical Longitudinal Data: Log-linear, panel, trend and cohort analysis**, London: Sage

Hand A., Gambles J. and Cooper E. (1994) **Individual Commitment to Learning. Individuals decision-making about lifetime learning**, Employment Department

Hume D. (1962) **On Human Nature and the Understanding**, New York: Collier

Jackson D. and Borgatta E. (1981) **Factor Analysis and Measurement**, London: Sage

Kim J. and Mueller C. (1978_a) **Introduction to Factor Analysis. What it is and how to do it**, London: Sage

Kim J. and Mueller C. (1978_b) **Factor Analysis: Statistical methods and practical issues**, London: Sage

Kruskal J. and Wish M. (1978) **Multidimensional Scaling**, London: Sage

Lee E., Forthfer R. and Lorimar R. (1989) **Analyzing Complex Survey Data**, Sage: London

Lehtonen R. and Pahkinen E. (1995) **Practical methods for design and analysis of complex surveys**, Chichester: John Wiley and Sons

Main B. and Shelly M. (1990) The effectiveness of the Youth Training Scheme as a manpower policy, in **Economica**, 57, pp. 495-514

Mare R. and Winship C. (1985) School Enrollment, Military Enlistment, and the Transition to Work: Implications for the age pattern of employment, p. 364 in Heckman J. and Singer B. (Eds.) **Longitudinal Analysis of Labor Market Data**, London: Cambridge University Press

Marradi A. (1981) Factor Analysis as an Aid in the Formation and Refinement of Empirically Useful Concepts, p. 11, in Jackson D. and Borgatta E. (Eds.) **Factor Analysis and Measurement**, , London: Sage

Martin J. and Roberts C. (1984) **The women and employment survey: a lifetime perspective**, London: HMSO

Maxwell A. (1977) **Multivariate Analysis in Behavioural Research**, New York: Chapman and Hall

McNabb R. and Whitfield K. (1994) **The Market for Training**, Aldershot: Avebury

Norusis M. (1994) **SPSS Advanced Statistics 6.1**, SPSS Inc.

Park A. (1994) **Individual Commitment to Lifelong Learning: Individuals attitudes. Report on the quantitative survey**, Employment Department

Park A. and Tremlett N. (1995) **Individual Commitment to Learning. Comparative findings from the surveys of Individuals, Employers and Providers attitudes**, Employment Department

Pedhazur E. (1982) **Multiple Regression in Behavioural Research**, London: Holt, Rhinehart and Winston

Pettigrew A., Hendry C. and Sparrow P. (1989) **Training in Britain. A study of funding, activity and attitudes. Employers perspectives on human resources**, London: HMSO

Pifer L. and Miller J. (1995) **The Accuracy of Student and Parent Reports About Each Other**, paper presented at American Educational Research Association Annual Conference, San Francisco

Plewis I. (1990) Longitudinal Multilevel Models, p. 119 in Dale A. and Davies R. (Eds.) **Analyzing Social and Political Change: A casebook of methods**, London: Sage

Reynolds H. (1977) **Analysis of Nominal Data**, London: Sage

Roberts K. and Parsell G. (1990) **Young People's Routes into UK Labour Markets in the Late 1980s**, ESRC 16-19 Initiative occasional papers 27, London: City University

Roker D. (1991) **Gaining the Edge: The education, training and employment of young people in private school**, London: City University

SCELI (1991) **Social Change and Economic Life Initiative Surveys 1986-1987**, Colchester: ESRC Data Archive

Schiffman S., Reynolds M. and Young F. (1981) **Introduction to multidimensional scaling. Theory, methods and applications**, London: Academic Press

Siegel S. (1956) **Nonparametric Statistics**, Tokyo: McGraw-Hill

Stevens J. (1992) **Applied Multivariate Statistics for the Social Sciences**, London: Lawrence Erlbaum

Taylor S. and Spencer L. (1994) **Individual Commitment to Lifelong Learning: Individual's attitudes. Report on the qualitative phase**, Employment Department

Tuma N. (1994) Event History Analysis: An introduction, p. 138 in Dale A. and Davies R. (Eds.) **Analyzing Social and Political Change: A casebook of methods**, London: Sage

Welsh Office (1995) **1994/95 Welsh Training and Education Survey**, Cardiff: Welsh Office

Whitfield K. and Bourlakis C. (1991) An empirical analysis of YTS, employment and earnings, in **Journal of Economic Studies**, 18, 1, pp. 42-56

Wisdom J. (1987) **Philosophy of the Social Sciences 1: A metascientific introduction**, Aldershot: Avebury

APPENDIX Outline file structure and list of variables

INITIAL EDUCATION

id	case
Area	research site
type	training trajectory
a1yr	year of birth
a1age	age now
a2	type of school
a3	regular attender
a4	attempted qualifications at school
a05	qualifications at school
a6a	continued with FE
a6b	reason for leaving school
a6num	number of FE/HE episodes
a7.1age - a7.3age	age at period 1 to 3
a7.1yr - a7.3yr	year of period 1 to 3
a8.1len - a8.3len	length of period 1 to 3 in months
a9.1 - a9.3	type of school period 1 to 3
a10.1wb - a10.3wb	course type 1 to 3
a11.1 - a11.3	reason for starting period 1 to 3
a12a.1 - a12a.3	attempted qualifications period 1 to 3
a12b.1 - a12b.3	qualification period 1 to 3
a13.1 - a13.3	reason for ending period 1 to 3
a14	age of leaving FTCE
a15	why left FTCE at this stage
a16	highest qualification at end of FTCE
a17	highest lifetime qualification
a18	most prestigious occupational class
a19	modal occupational class
a20	number of lifetime learning episodes
a21	number of lifetime training episodes
a22	years spent in lifetime learning
a23	proportion of lifetime in learning

WORK HISTORY

b0num	life history index (MULTIPLE RECORDS PER CASE)
b1	economic status
b2age	age at start
b2yr	year of start
b3len	length in months
b4wb	industrial classification
b5	public/private
b6a	occupational class
b6b	social class
b7	number of people supervised
b8	how many people in company
b9	availability of training a factor
b10	any training received
b11aa	any first aid etc
b11ab	any one-day IT etc
b11b	any other training
b12a	induction training
b12amr	length of induction
b12b	on job training
b12bmr	length of on job
b12c	new task training
b12cmr	length of new task
b12d	transferable training
b12dmr	length of transferable
b13num	number of substantive training episodes per post
b13.1age - b13.4age	age at start of training period 1 to 4
b13.1yr - b13.4yr	year of start of training period 1 to 4
b14.1len - b14.4len	length in years of period 1 to 4 (weeks)
b15.1 - b15.4	hours per week period 1 to 4
b16.1- b16.4	during normal working hours period 1 to 4
b17.1- b17.4	who paid period 1 to 4
b18.1- b18.4	provider period 1 to 4
b19.1 - b19.4	course type 1 to 4
b20.1 - b20.4	reason for starting period 1 to 4
b21a.1 - b21a.4	attempted qualifications period 1 to 4
b21b.1 - b21b.4	qualification period 1 to 4
b22.1- b22.4	reason for ending period 1 to 4
b23.1- b23.4	lead to promotion period 1 to 4

POST-CONTINUOUS EDUCATION

c1	any other education
c1num	number of POCE episodes
c2.1age - c2.4age	age at start of period 1 to 4
c2.1yr - c2.4yr	year of start of period 1 to 4
c3.1len - c3.4len	length of period 1 to 4 (weeks)
c4.1 - c4.4	hours per week period 1 to 4
c5.1- c5.4	organiser period 1 to 4
c6.1 - c6.4	who paid 1 to 4
c7.1 - c7.4	course type 1 to 4
c8.1 - c8.4	reason for starting period 1 to 4
c9a.1 - c9a.4	attempted qualifications period 1 to 4
c9b.1 - c9b.4	qualification period 1 to 4
c10.1 - c10.4	reason for ending period 1 to 4

c12a	any hobbies
c12b.1 - c12b.2	type of hobbies 1 to 2
c12c.1age - c12c.2age	age at start of hobbies 1 to 2
c12c.1yr - c12c.2yr	year of start of hobbies 1 to 2
c12d.1age - c12d.2age	age at end of hobbies 1 to 2
c12d.1yr - c12d.2yr	year of end of hobbies 1 to 2

FAMILY BACKGROUND

d1	ethnicity
d2	family religion
d3	language
d4	marital status
d5age	age of leaving parents home
d5yr	year left parents
d6a.1age - d6a.3age	age start partnerless period 1 to 3
d6b.1age - d6b.3age	age end partnerless period 1 to 3
d6a.1yr - d6a.3yr	year start partnerless 1 to 3
d6b.1yr - d6b.3yr	year end partnerless 1 to 3
d7	place of birth
d8num	number of moves
d9.1age - d9.7age	age of area move 1 to 7
d9.1yr - d9.7yr	year of area move 1 to 7
d9.1 - d9.7	area moved to 1 to 7
d10	length of residential association with research site
d11	length of residential association with South Wales

e1.1yr - e1.4yr	year born, partner, father, mother, sibling
e1gende	gender sibling
e2.1age - e2.4age	age left school, partner, father, mother, sibling
e3.1 - e3.4	qualification, partner, father, mother, sibling
e4.1a - e4.4a	occupational classification, partner, father, mother, sibling
e4.1b - e4.4b	social class, partner, father, mother, sibling
e5.1 - e5.4	economic status, partner, father, mother, sibling
e6.1 - e6.4	area, partner, father, mother, sibling

f1	any children
f1num	number of children
f2.1age - f2.4age	age of giving birth child 1 to 4
f2.1yr - f2.4yr	year child 1 to 4

f3.1yr - f3.4yr	year child 1 to 4 left home
f3.1age - f3.4age	age child 1 to 4 left home
f4.1age - f4.4age	age child 1 to 4
f5.1 - f5.4	type of school child 1 to 4
f6.1 - f6.4	gender child 1 to 4
f7.1 - f7.4	regular attender child 1 to 4
f8.1 - f8.4	left FTCE child 1 to 4
f8.1age - f8.4age	age at end of FTCE child 1 to 4
f9.1 - f9.4	qualification child 1 to 4
f10.1a - f10.4a	occupational status child 1 to 4
f10.1b - f10.4b	social class child 1 to 4
f11.1 - f11.4	economic status child 1 to 4
f12.1 - f12.4	living at home child 1 to 4
f13.1 - f13.4	area of residence child 1 to 4

g1	respondents gender
g2	age cohort
g5	willing to be interviewed

ISBN 1 872330 03 7

99